MULTI-SCALE MOVEMENT OF DEMERSAL FISHES IN ALASKA

By

Julie K. Nielsen, B.A., B.S., M.S.

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Fisheries

University of Alaska Fairbanks

May 2019

APPROVED:

Andrew C. Seitz, Committee Chair
Timothy Loher, Committee Member
Susanne F. McDermott, Committee Member
Franz J. Mueter, Committee Member
Milo D. Adkison, Committee Member and Chair
Department of Fisheries
S. Bradley Moran, Dean
College of Fisheries and Ocean Sciences
Michael Castellini, *Dean of the Graduate School*

Abstract

Information on the movement of migratory demersal fishes such as Pacific halibut, Pacific cod, and sablefish is needed for management of these valuable fisheries in Alaska, yet available methods such as conventional tagging are too coarse to provide detailed information on migration characteristics. In this dissertation, I present methods for characterizing seasonal and annual demersal fish movement at multiple scales in space and time using electronic archival and acoustic tags. In Chapter 1, acoustic telemetry and the Net Squared Displacement statistic were used to identify and characterize small-scale movement of adult female Pacific halibut during summer foraging in a Marine Protected Area (MPA). The dominant movement pattern was home range behavior at spatial scales of less than 1 km, but a more dispersive behavioral state was also observed. In Chapter 2, Pop-up Satellite Archival Tags (PSATs) and acoustic tags were deployed on adult female Pacific halibut to determine annual movement patterns relative to MPA boundaries. Based on observations of summer home range behavior, high rates of year-round MPA residency, migration timing that largely coincided with winter commercial fisheries closures, and the demonstrated ability of migratory fish to return to previously occupied summer foraging areas, the MPA is likely to be effective for protecting both resident and migrant Pacific halibut brood stock year-round. In Chapter 3, I adapted a Hidden Markov Model (HMM) originally developed for geolocation of Atlantic cod in the North Sea for use on demersal fishes in Alaska, where maximum daily depth is the most informative and reliable geolocation variable. Because depth is considerably more heterogeneous in many regions of Alaska compared to the North Sea, I used simulated trajectories to determine that the degree of bathymetry heterogeneity affected model performance for different combinations of likelihood specification methods and model grid sizes. In Chapter 4, I added a new geolocation variable, geomagnetic data, to the HMM in a small-scale case study. The results suggest that the addition of geomagnetic data could increase model performance over depth alone, but more research is needed to continue validation of the method over larger areas in Alaska. In general, the HMM is a flexible tool for characterizing movement at multiple spatial scales and its use is likely to enrich our knowledge about migratory demersal fish movement in Alaska. The methods developed in this dissertation can provide

valuable insights into demersal fish spatial dynamics that will benefit fisheries management activities such as stock delineation, stock assessment, and design of space-time closures.

Table of Contents

Page

List of Figures

Page

List of Tables

Page

Acknowledgments

My first words of thanks go to my advisor and committee members. I don't know if it is possible to thank Andy Seitz enough for the support, encouragement, ideas, and editing skills he has provided over the years. He has been an amazing advisor, and I look forward to collaborating with him on projects beyond this dissertation. I was also very lucky to have Milo Adkison, Tim Loher, Susanne McDermott, and Franz Mueter on my committee. Together they provided a wide range of expertise in modeling, statistics, biology, and management that added a great deal of value to my research; I hope to collaborate with them on future research projects as well.

Second, I would like to thank the funding sources that made this research possible. I received fellowships from the Rasmuson Fisheries Research Center and the Pollock Conservation Cooperative Research Center that funded the majority of my studies. It was a pleasure to meet and interact with the advisory board members of these research centers, and I will do my best to provide a return on their investment by continuing to conduct research in support of fisheries management in Alaska. I also received funding from the National Park Service to conduct Pacific halibut research in Glacier Bay (Chapters 1 and 2); I would like to thank Chad Soiseth for helping to obtain this funding as well as assistance with the research. The North Pacific Research Board also contributed to my research through a graduate research fellowship. These funds made it possible to conduct the acoustic tracking for Chapter 2, which added immense value to the project. Mark Evans and the UAF Undergraduate Research & Scholarly Activity program provided the acoustic tags for the acoustic tracking for Chapter 2.

Third, I would like to thank specific people who directly supported my research. Jim Taggart and Philip Hooge collected the Pacific halibut acoustic data that were analyzed in Chapter 1 and provided valuable insights into animal movement and behavior. For help with fieldwork in Chapter 2, I would like to thank Zach Stenson (F/V Taurus), Jim de La Bruere and crew of R/V Medeia, Julien Appignanii, Anne Beaudreau, Amanda Compton, Michael Courtney, Mark Evans, Thomas Farrugia, Jessica Glass, Gordon Kruse, Tim Loher, Craig Murdoch, Katy Rayfield, John Rodstrom, Kevin Siwicke, Chad Soiseth, Alan Steffert, Barbara Stone, David Stone, and Maurice Tivey. Bob Stone provided a time series of

temperature measurements from stationary moorings in the Gulf of Alaska that greatly benefited the analyses for Chapter 2. For help with Chapter 3, I thank Martin Pedersen and Uffe Thygesen, who provided code for the HMM that was adapted for demersal fish in Alaska and advice about how to run the model. I also thank Arnault le Bris for additional advice and feedback on the HMM and Margaret Short for discussions about fish geolocation models in general. For Chapter 4, I thank David Stone and Maurice Tivey for their instruction on the Earth's magnetic field and instruments that can be used to measure it. I would also like to thank Katy Echave for supplying tags that provided *in situ* stationary data near Kodiak and discussions about the potential importance of geomagnetic data for sablefish.

Finally, I would like to thank people who supported me personally during this endeavor. Gabrielle Hazelton, Chris Brooks, and Lawrence Powers provided extensive academic and fieldwork support. The other students in the Seitz lab (Thomas Farrugia, Michael Courtney, Stephanie Meggers, Marguerite Tibbles, Kaitlyn Manishin, Parker Bradley, Kevin Siwicke, Matt Smukall, and Mark Evans) were a source of support and advice. Discussions at the Ted Stevens Marine Research Institute with Cindy Tribuzio, Katy Echave, Kalei Shotwell, Karson Coutre, Dana Hanselman, and Jon Heifetz about electronic tagging in general have been immensely helpful. The Schaefer family generously provided a place for me to stay during Fairbanks visits. I would like to thank my parents, Sally and Jim Nielsen, for a lifetime of support and encouragement. And last, but certainly not least, I would like to thank my husband, Alan Steffert, for putting up with the late nights, providing engineering advice, serving as a liaison with tag engineers, helping me with the fieldwork, and just being there for me – I am deeply grateful for all of that.

General introduction

Demersal fishes such as Pacific halibut (*Hippoglossus stenolepis*), Pacific cod (*Gadus macrocephalus*), and sablefish (*Anoplopoma fimbria*) are commercially, culturally, and ecologically valuable in Alaska. All are large-bodied benthic predators that can inhabit deep waters and are known to conduct large-scale seasonal (Shimada and Kimura, 1994; Loher and Seitz, 2006) or ontogenetic (Skud, 1977; Heifetz and Fujioka, 1991) migrations. However, most of the information on long-distance movements for these species has been obtained from conventional tag release and recovery locations alone, which cannot provide a detailed understanding of migration timing, pathways, spawning locations, or the proportion of the population that migrates. Thus, tagging methods that can provide detailed information about fish movements are needed to answer important questions about stock structure, movement between management areas, and seasonal migration patterns to support management of these species (Thompson and Dorn, 2004; Maloney and Sigler, 2008; Loher, 2011).

Electronic tags provide an alternative to conventional tags for gathering information on movements by allowing the reconstruction of movement trajectories for individual tagged fish. In general, electronic tags contain a microprocessor and a suite of sensors that measure and record information about environmental conditions such as ambient light intensity, depth, and temperature. Tags vary by model and manufacturer in the type and number of sensors they possess, the format of the data collected, and the frequency of measurements recorded by the tag. The tags are designed either to be recovered when the fish is recaptured ("harvest-recovered tag") or to release from the fish at a specified time and broadcast summarized daily data to a satellite network ("Pop-up Satellite Archival Tags", or PSATs). Data transmitted by PSATs are typically summarized by 12- or 24-hour periods. However, some PSATs are capable of fully uploading detailed archival data with a temporal resolution that is inversely proportional to the duration of the tag deployment. In conjunction with release and recovery locations, electronic tag data can be used to reconstruct daily locations of the tagged fish as long as 1) the tags measure environmental variables that vary over space in the study area, 2) those environmental variables are accurately mapped in the study area, and 3) tag and map resolution are adequate for detecting variation

within the range of movement of the tagged animal in the study area. If the above conditions are met, estimated locations of tagged animals can then be generated by matching values measured by the tags to maps of geolocation variables in the study area. This process is referred to as "geolocation."

Methods for geolocation vary based on the types of electronic tags used and availability of maps and spatial gradients of environmental values in the study area. Variables such as depth or sea surface temperature (SST) can be measured by satellites or vessel surveys and are available in discrete grid formats. Other variables are "mapped" through continuous spatial models. For example, ambient light intensity measurements which vary spatially and seasonally can provide daily estimates of latitude and longitude based on day length and the time of local noon recorded by the tag each day (Musyl et al., 2001). For pelagic fishes in lower latitudes, such as tunas and billfishes, light and SST are the primary environmental variables used for geolocation (Lam et al., 2008; Schaefer et al., 2011). However, these variables are not appropriate for demersal fish species in high-latitude locations because light intensity data alone at high latitudes does not provide many locations, especially when fishes occupy depths below 150 m (Seitz et al., 2006). Some demersal fish species such as sablefish occupy waters that are too deep to detect any light at all. In addition, even if light is adequate for producing position estimates, temperature cannot be used to filter position estimates because bottom temperatures differ from SST. Thus, alternatives to light- and SST-based geolocation are needed for reconstructing movement trajectories of demersal fish species in Alaska.

Sophisticated state-space movement models can be used to reconstruct movement pathways of tagged demersal fishes using the geolocation data recorded by electronic tags (Lam et al., 2008; Patterson et al., 2008; Pedersen et al., 2008; Jonsen et al., 2013). These models consist of an underlying movement model, which contains information about how the tagged fish is expected to move through the study area, and a data likelihood model that determines how well the geolocation data collected by the fish matches the mapped values in the study area at different locations. By explicitly allowing the inclusion of prior information on fish movement speed and behavior, these models can integrate information about movement at small scales in space and time (such as daily movement rates) to inform estimates of large-

scale movement. Therefore, in the context of movement states for the state-space geolocation model, knowledge of small-scale behavior such as daily movement patterns and habitat associations is useful for a comprehensive understanding of movement at annual timescales.

The goal of this dissertation is to develop methods for characterizing the movements of migratory demersal fish species in Alaska at multiple scales using two types of electronic tags and a state-space geolocation model. Acoustic tags are used to determine fine-scale movement patterns and movement speeds during the summer foraging period. Archival tags (PSATs) are used to determine movement patterns over annual timescales. A Hidden Markov Model (HMM) is adapted to reconstruct movement of demersal fishes in Alaska based on archival and acoustic tag data. In general, Pacific halibut is the focal species of this dissertation, but the methodology developed and presented here can be applied to any other demersal species in Alaska.

Chapter one provides information on fine-scale movement patterns and scales of movement for adult Pacific halibut during summer foraging in Glacier Bay National Park, Alaska based on acoustic telemetry. In addition to describing important insights into site fidelity and habitat associations for Pacific halibut, the analysis method based on Net Squared Displacement (NSD) employed in this chapter provides a new way to characterize fish movement based on quantifying dispersion over time rather than analyzing spatial patterns (such as kernel densities) alone. This method is robust to small sample sizes and irregular data sets, and provides results in the form of potential for dispersion, which is the format needed for large-scale movement modeling applications.

Chapter two extends the information on seasonal movement patterns of adult Pacific halibut described in Chapter one to provide insights into annual movement patterns relative to Marine Protected Area (MPA) boundaries using both acoustic and archival tags. A combination of PSATs and acoustic transmitters is used to infer whether tagged fish are inside or outside of Glacier Bay for each day the tagged animals are at liberty. This chapter provides important information on partial migration, site fidelity, and annual movement of adult female Pacific halibut relative to MPA boundaries and the timing of the commercial fishery. The results from this chapter combine knowledge of movement at different

spatial scales to provide insights into the potential effectiveness of the Glacier Bay MPA for protecting Pacific halibut brood stock year-round.

Chapter three focuses on methods for reconstructing movement paths of demersal fishes over annual timescales using archival tag data with a state-space geolocation model. An HMM developed for Atlantic cod in the North Sea (Pedersen et al., 2008) was adapted to accommodate the different types of data and environmental gradients that occur in Alaska. The data likelihood model is based primarily on depth, the most informative geolocation variable in Alaska, although other sources of geolocation data may be included when available. In this chapter, the effect of different degrees of bathymetric heterogeneity present in three regions of Alaska is explored. Estimated locations from simulated trajectories are used to provide insights into optimal data likelihood specification methods and model grid sizes for demersal fishes in different regions of Alaska. Though focused primarily on depth, this chapter provides important insights into selecting optimal likelihood specifications and grid sizes for other gridded geolocation variables such as SST and geomagnetic data.

Chapter four extends the methodology described in Chapter three by adding geomagnetic data to the data likelihood model for the HMM. This chapter provides basic information on the Earth's magnetic field, its variation in space and time, and how geomagnetic data can be used for geolocation in a high-latitude magnetic anomaly area. Estimated locations from simulated trajectories are used to determine whether data likelihood models based on magnetic anomaly maps can improve geolocation over the use of depth alone and the resolution of geomagnetic tags required for improvement to be observed. However, further validation is needed to ensure map accuracy and tag resolution are adequate in all regions of Alaska.

Taken together, these four chapters demonstrate the ways that electronic tags can provide more detailed information on annual movement patterns than tags that simply provide release and recovery locations alone. For Pacific halibut, these methods have resulted in new insights into behaviors such as site fidelity that are important considerations for management of the fishery such as spatial and temporal

closures of preferred halibut fishing areas. Important insights into annual movement patterns of Pacific cod and sablefish may also be revealed using methods developed in this dissertation.

Chapter 1. Characterizing Pacific halibut movement and habitat in a Marine Protected Area using net squared displacement analysis methods[1]

1.1 Abstract

We characterized small-scale movement patterns and habitat of acoustic-tagged adult (68 – 220 cm TL) female Pacific halibut during summer and fall in Glacier Bay National Park, Alaska, a marine protected area (MPA). We used net squared displacement (NSD) analysis methods to identify two movement states, characterize individual dispersal patterns, and relate habitat variables to movement scales. Movement states identified for 32 of 43 halibut consisted of 1) a non-dispersive "residential" movement state (n = 27 fish), where movement was restricted to an average movement radius of 401.3 m (95% C.I. 312.2 – 515.9 m) over a median observation period of 58 days, and 2) a "dispersive" movement state (n = 15 fish), where movements of up to 18 km occurred over a median observation period of 27 days. Some fish (n = 10) exhibited both movement states. Individual fish demonstrated primarily non-random dispersal patterns including home range (n = 17), site fidelity (return to previously occupied locations following forays, n = 6), and shifted home ranges (n = 5). However, a random dispersal pattern (n = 4) with an estimated diffusion rate of 0.9 ± 0.05 km^2 day^{-1} SE was also observed. Home range size increased with depth but not fish size. Home range locations were associated with heterogeneous habitat, intermediate tidal velocities, and depths < 100 m. Observations of non-dispersive movement patterns, relatively small home ranges, and site fidelity for adult females suggest that MPAs such as Glacier Bay may have utility for conservation of Pacific halibut broodstock.

1.2 Introduction

Knowledge of fish movement patterns at multiple spatial and temporal scales can benefit the management of mobile fish species. For example, even highly migratory species that move thousands of

[1] This is a pre-copyedited, author-produced version of an article accepted for publication in Marine Ecology Progress Series following peer review. The version of record [Nielsen, J. K., Hooge, P. N., Taggart, S. J., and Seitz, A. C. 2014. Characterizing Pacific halibut movement and habitat in a Marine Protected Area using net squared displacement analysis methods. Marine Ecology Progress Series, 517: 229-250] is available online at: https://doi.org/10.3354/meps11043.

kilometers are capable of philopatry at very small spatial scales (Jorgensen et al. 2010). Thus, understanding small scale movement patterns and habitat associations can be an important part of achieving a wholistic understanding of stock dynamics and assessing the potential effectiveness of spatial management techniques such as marine protected areas (MPAs) for migratory fish species.

The Pacific halibut, *Hippoglossus stenolepis* (hereafter referred to as 'halibut'), is an economically, ecologically, and culturally important flatfish species in the North Pacific Ocean. Based on observations of large-scale seasonal and ontogenetic movements during larval, juvenile, and adult life history stages (Valero & Webster 2012), halibut in North America are managed on a large scale (Clark & Hare 2006) where a single stock assessment is conducted for a region that ranges from California to the Bering Sea before the allowable harvest is apportioned into smaller management units (Webster & Stewart 2014). Some proportion of adult halibut conduct seasonal spawning migrations from summer foraging locations in near-shore areas to winter off-shore spawning areas in deeper waters on the continental slope of the Pacific Ocean (Loher & Seitz 2006, Loher 2011, Seitz et al. 2011). Recent pop-up satellite archival tagging and conventional tagging research has demonstrated that a large proportion of adult halibut exhibit inter-annual site fidelity and homing to summer foraging locations (Loher 2008). These observations suggest that knowledge of movement patterns at smaller scales will be important for understanding the spatial sub-structure of the halibut stock, potential local effects of intense fishing, and the utility or effectiveness of MPAs as a management tool for halibut.

In addition to coastwide, large-scale management through area-specific harvest rates, halibut are also regulated at smaller spatial scales through catch sharing plans as well as the existence of MPAs. For example, halibut harvest is restricted in the interior waters of Glacier Bay National Park in southeastern Alaska, where commercial fishing for halibut is being phased out (36 CFR 13.1130-1146) over several decades and sport fishing is limited by daily vessel quotas (36 CFR 13.1150-1160) during the summer months. Glacier Bay National Park was added to the National System of Marine Protected Areas in 2009. As a large, high-latitude MPA, Glacier Bay may eventually protect halibut that reside within its boundaries from commercial harvest. However, obtaining information on the scale and patterns of halibut

movement and habitat associations is critical for understanding Glacier Bay's potential effectiveness at retention of adults (Kramer & Chapman 1999) and specific benefits that may result from protection.

Here, we present information on the spatial and temporal scales of movement by adult halibut in Glacier Bay National Park during summer and fall that may be valuable for assessing the potential effectiveness of Glacier Bay National Park as an MPA. We use net squared displacement (NSD) analysis techniques to 1) identify and characterize two distinct movement states, "residential" and "dispersive", 2) classify and quantitatively describe dispersal patterns for individual tagged halibut, and 3) describe habitat associations and relationships between habitat variables (depth, average tidal speed, habitat complexity, and substrate type) and scale of movement for the residential movement state. We interpret these results in terms of spatially explicit fisheries management applications such as MPA design and effectiveness. We conclude by addressing the potential contribution of NSD analysis methods for characterizing the movement patterns and dispersal scales of fishes and facilitating MPA design.

1.3 Materials and methods

1.3.1 Study area

The study was conducted in the northern portion of southeastern Alaska within the inside waters of Glacier Bay National Park (Figure 1.1). Glacier Bay is a glacial fjord that is influenced by both current and historical glacial activity. Glaciers have receded more than 100 km in the last 300 years in Glacier Bay, leaving behind a Y-shaped body of water with deep (200 – 450 m) marine basins interspersed with shallow moraines and tidewater glaciers at the heads of the fjords. Substantial glacial freshwater runoff influences the oceanography with high sedimentation and areas of cold water upwelling. Strong tidal currents mix the water column completely in the shallow lower portion of the bay, but deeper upper reaches are largely stratified. Primary productivity levels are highest in a transition zone in the central portion of the bay that is characterized by intermediate stratification. Salinity, temperature, and light penetration decrease towards the heads of the fjords (Etherington et al. 2007b).

1.3.2 Fish tagging and tracking

A total of 43 halibut were captured on longlines, tagged and released in Glacier Bay during the

summers of 1991–1993 (Table 1.1, Fig. 1.1). Longlines were set at four general release locations within the study area using snap-on gangions designed for the commercial halibut fishery and were "soaked" for 6 hours. Capture locations were determined when each fish was brought on board the capture vessel using a PLGR GPS that removed selective availability errors. We generally selected larger fish (> 100 cm) for tagging because we were primarily interested in fish that were vulnerable to the commercial fishery (≥ 82 cm) and we wanted to minimize possible effects of large, long-life acoustic tags on behavior.

Acoustic transmitters (Sonotronics, Tucson, Arizona) that transmitted a unique identifying sonic pulse were attached to halibut externally during 1991 and 1992 (n = 26) and internally during 1992 and 1993 (n = 17). Externally attached acoustic tags were secured to fish by inserting two Teflon-coated stainless steel wires through the dorsal musculature immediately ventral to the dorsal fin, with a backing plate of neoprene rubber and fiberglass. A sterilized needle was used to thread the wire. For the internal attachment, tags were surgically implanted in the coelomic cavity using sterile methods. Tags were inserted into the coelomic cavity through a 5-cm incision on the eyed-side, parallel and 2 – 3 cm dorsal to the long axis of the fish. Seven to eight external sutures (2-0 Braunamid non-absorbable) were used to close the incision. During the 5 – 15 minute surgery, the gills of the fish were irrigated with ambient seawater which was well-mixed and high-saline in the study area. When possible, information on the sex of the tagged animal was obtained through cannulation or observation during surgical implantation. Tagged animals were released within 500 m of the location where they were brought on board.

Acoustic tag transmission frequency and size varied during the study. Acoustic tags attached during the first year (n = 9) transmitted at a frequency of 80 kHz, whereas 35 kHz acoustic tags (n = 34) were used in the two subsequent years due to their increased detectability in Glacier Bay's waters. We used two sizes of acoustic tags in the study. The smaller tags (n = 17) were 95 mm long x 18 mm diameter, weighed 16 g in water, and had an observed lifetime of 1.3 – 2 years. The larger tags (n = 26) were 95 mm long x 34 mm diameter, weighed 34 g in water, and had an observed lifetime of 2.5 – 3.4 years. Details of tag attributes for individual tagged animals are provided in Supplement 1-1 (Table S1.1-1).

Tagged halibut were tracked from a vessel using a bow-mounted dual hydrophone assembly lowered 2 m beneath the surface of the water and capable of rotating 360°. One hydrophone faced forward and -10° from horizontal and the other hydrophone pointed downward. These directional hydrophones (Sonotronics DH-2) had a beam width of ±6° and a sensitivity of 84 dBV and were connected to manual receivers (Sonotronics USR-4D). With this configuration, *in situ* range tests indicated that tags could be detected at distances of up to 2 km. When a tag was detected, the vessel operator maneuvered the vessel in a circular pattern in the vicinity of the tag until signal strength was uniform at all points on the circle and the signal received on the downward-facing hydrophone in the middle of the circle was highly amplified. A GPS was used to obtain the location of the vessel at this position, which served as the estimated position of the tagged animal. Positions of tagged animals were obtained daily to weekly during tracking periods that lasted 3 – 6 months, mostly in the summer and fall, of each year. Searches for tagged animals were conducted in an outward spiral starting from each individual's last known position. Consequently, if a tagged halibut moved more than a few kilometers away, it was not necessarily found during the subsequent search. An example of the spatial distribution of tracking effort (number of days tracked per season) in the study area during 1991 is shown in Supplement 1-1 (Figure S1.1-2).

The precision of position estimates for tagged animals was likely to decrease with increasing water depth. We estimated the precision of each observation based on a linear regression of error radii vs. depth for 1) known positions of tags recovered by SCUBA divers (n = 3) and 2) root mean squared distances between repeated observations of motionless tags (n = 6). The depth of each observation was multiplied by the resulting slope coefficient, 0.65 ($r^2 = 0.83$, $P = 0.0005$), to obtain error buffers for each observation that ranged from approximately ±10 m at depths of 10 m to approximately ±100 m at depths of 150 m.

1.3.3 Data analysis

Due to the large study area and the opportunistic nature of the animal resightings, the dataset was characterized by irregular sampling intervals, unequal sample sizes among animals, and small numbers of

observations for some animals. Because most movement analysis methods require regular and frequent observations of tagged animals, we employed an alternative analysis framework that is robust to missing data and small sample sizes. This analysis framework, based on the Net Squared Displacement (NSD) statistic, is based on the identification of patterns of dispersal over time that correspond to different behaviors such as foraging or migration (Börger & Fryxell 2012). NSD, also commonly referred to as R^2_n, is the square of the distance between the origin of a given trajectory and each subsequent position.

$$NSD_t = \left| x_t - x_0 \right|^2 \quad \text{(eq 1.1)}$$

where $\mathbf{x}_t$ is the coordinate vector at time t (i.e., the latitude and longitude of a fish on day t), and $\mathbf{x}_0$ is the coordinate vector for the origin of the trajectory (i.e., a fish's release location). For random movement, e.g. Brownian motion, the NSD statistic increases linearly with time (Kareiva & Shigesada 1983) and the slope is proportional to the rate of diffusion (Börger & Fryxell 2012). For non-dispersive movement, such as home range behavior, the NSD statistic reaches a constant value over time that represents the spatial scale of the area in which the animal moves (Turchin 1998, Moorcroft & Lewis 2006). For directed movement toward a specific location, such as during migration or moving between foraging locations, the relationship between NSD and time is exponential (Nouvellet et al. 2009).

1.3.4 Movement states

We defined two different movement states using the NSD statistic. The first movement state, "residential," reflects non-dispersive movement and was defined when the slope of NSD vs. time = 0 ($p > 0.05$ for the slope coefficient in a linear regression) for a minimum sample size of 4 consecutive observations (Figure 1.2). For this movement state, the intercept of NSD vs. time provides information about the spatial scale at which NSD values do not increase or decrease over time, thus providing an estimate of home range size that is robust to small sample sizes and infrequent observations (Moorcroft & Lewis 2006). Consecutive observations classified as residential are subsequently referred to as "home ranges". We estimated typical home range size for the residential movement state using an intercept-only linear mixed-effects model:

$$\log(\text{NSD})_{ij} = \alpha + a_i + \varepsilon_{ij} \quad \text{(eq 1.2)}$$

$$a_i \sim N(0,\sigma^2_a),\ \varepsilon_{ij} \sim N(0, \sigma^2)$$

where α is the fixed-effects estimate of mean home range size for the population of *i* home ranges, a_i is a random variable that represents the variation of individual home range estimates around the fixed-effects mean, and within-group error ε_{ij} is assumed to be independent and normally distributed with mean zero. The model was fit using restricted maximum likelihood. NSD values were log-transformed to account for heteroscedasticity prior to modeling. We applied the bias-correction for a log-normal distribution to back-transform the estimated mean to the original scale:

$$\text{Home range size (NSD)} = \exp(\alpha + s^2/2) \quad \text{(eq 1.3)}$$

where α is the mean home range size for the population as described above and s^2 is the estimated variance for α. To provide a more intuitive linear description of movement scale, the square root of intercept coefficients were reported as a "home range radius." To minimize potential bias of capture and tagging on the scale of home range movement patterns (e.g., temporary tagging effects, uncertainty in longline position during capture vs. release location, or travel from release position back to the home range), observations within 3 days of tagging were not used for home range analyses.

The second movement state, "dispersive," was classified as all observations where the slope of NSD vs. time ≠ 0 (Figure 1.2C). This movement state generally contained observations from both random and directed movement types. Because animals with more mobile movement patterns were more difficult to relocate, observations were not collected frequently enough to determine whether individual observation sequences were random (linear relationship for NSD vs. time) or directed (exponential relationship for NSD vs. time). Therefore, we did not summarize this movement state using a mixed-effects model, as we did for the residential movement state, because it likely contained a mix of both movement types.

Although we were not able to explicitly classify observations as either random or directed, insight into the randomness of the dispersive movement state was obtained by comparing our observations of average NSD vs. time to expected random values using correlated random walks (CRWs). CRWs are

movement paths comprised of a discrete series of movement steps where the expected distance and the expected angle between subsequent steps determines overall movement path characteristics such as diffusion rates (larger for larger step lengths) and directed vs. random movement (more directed for smaller variation in turning angles, more random for larger variation in turning angles). CRWs are usually simulated based on empirical distributions of both step lengths and turning angles obtained from frequent and regular observations of an animal's location over time (Kareiva & Shigesada 1983, Turchin 1998). However, because our dataset was highly irregular, we simulated random movement from a step length distribution comprised of all records that were one day apart, but used a theoretical distribution of turning angles (wrapped Cauchy with an autocorrelation coefficient of 0.5) that is typically observed during animal foraging activities (Morales et al. 2004, Bartumeus et al. 2005). This approach allowed us to incorporate empirical information on spatial scales of our tagged animals (daily step lengths) under a specific hypothesis of random movement during foraging. We fit exponential curves (Moorcroft & Lewis 2006) to both the residential and dispersive step length distributions and sampled randomly from these distributions to create 1000 CRWs for a duration of 90 days for each movement state. To account for the effects of movement that occurs within the confined waters of the study area, simulations were conducted on a 20 m bathymetry grid of Glacier Bay (Geiselman et al. 1997) and were initiated at the first location of each observed residential or dispersive movement sequence. If a simulated position for the CRW fell on land, that position was discarded and a new coordinate was chosen. To assess randomness, the mean NSD from observed data was compared to the mean and 95% C.I. from the CRWs at each time step. NSD values for random movement should fall within the 95% C.I. for the CRW simulations, whereas values for directed and non-dispersive movement will be greater than the upper bound and less than the lower bound, respectively, of the 95% C.I. for CRW values (Austin et al. 2004).

1.3.5 Individual dispersal patterns

In addition to characterization of movement states, knowledge of the way in which NSD changes over time during non-dispersive, directed, or random movement can be used to formulate theoretical models that describe behavioral phenomena such as home range occupation, directed movements to new

locations, migrations or forays, or random movement (Bunnefeld et al. 2011, Börger & Fryxell 2012, Papworth et al. 2012, Singh et al. 2012). We identified five models of NSD vs. time, subsequently referred to as dispersal patterns, that we assume represent underlying behaviors for tagged animals in this study (Table 1.2). First, the "home range" dispersal pattern (HR) reflects a restricted range of movements described by a slope of zero for NSD vs. time. Second, the "random" dispersal pattern (R) reflects diffusion (e.g., Brownian motion) and can be described by a linear increase in NSD vs. time. Third, the "shifted home range" dispersal pattern (SHR) represents movement from the release location to another location in the study area and consists of a non-linear (sigmoidal) model with parameters for the timing of the midpoint of the travel to the new location (θ, Julian day), a scale parameter (φ, days) to estimate the time to travel between the midpoint and approximately ¾ of the distance to the destination (Bunnefeld et al. 2011), and the squared distance to the destination (δ, m^2). Fourth, the "site fidelity" (SF) dispersal pattern represents departure from the release location and subsequent return to the original location. It consists of a double sigmoidal model, with one sigmoid function to describe the migration start (subscript m) and one to describe the return to the original location (subscript r). The site fidelity dispersal pattern has the same parameters as the shifted home range model, but with an additional parameter, θ_r, to describe the timing (Julian day) of the midpoint for the return. Finally, the "foray/shifted home range" dispersal pattern (FSHR) reflects departure from the release location and dispersive behavior before resuming home range behavior in another location. The parameters for this model are the same as those for the site fidelity pattern, but an additional parameter (δ_r, m^2) is added to describe the squared distance between the farthest distance traveled during the foray and the shifted home range.

Dispersal patterns for individual tagged animals were classified by determining which of the five models for NSD vs. time described above best described the observed values of NSD vs. time based on model selection techniques. Non-linear models (all models besides home range and random) were fit using a least squares algorithm (nls) as implemented in the "nlme" package for R (Pinheiro & Bates 2000). The best-fitting model for each individual trajectory was selected using the Akaike Information Criterion adjusted for small sample sizes (AICc) (Burnham & Anderson 1990) and residual analysis. An example

of the dispersal pattern classification process for individual animals is provided in Supplement 1-2 (Figure S1.2-1). Once individual fish were classified according to dispersal pattern, we calculated 1) the average maximum distance from the release location during the observation period, 2) the average distance from the release location at the end of the observation period, 3) average observation period duration, and 4) average animal size for each dispersal pattern.

We used mixed-effects models to summarize model parameters for dispersal patterns to which more than 3 animals were assigned. Because animals with the home range dispersal pattern were included in the mixed-effects model for the residential movement state, mixed-effects models were only used to summarize the random and site fidelity dispersal patterns. For the random dispersal pattern, we quantified the rate of dispersal over time using a linear model with no intercept, as by definition dispersal must be zero at the origin of a trajectory, using animal ID as a grouping variable.

$$NSD_{ij} = \beta t + b_i t + \varepsilon_{ij} \quad \text{(eq 1.4)}$$

$$b_i \sim N(0,\sigma^2_b),\ \varepsilon_{ij} \sim N(0, \sigma^2)$$

where β is the fixed-effect variable estimate of the mean slope of NSD vs. time for the population of i random trajectories, b_i is a random variable that represents the variation of individual slopes around the population mean slope, and within-group errors ε_{ij} are independent and normally distributed with mean 0. The model was fit using restricted maximum likelihood. Because random movement results in the process of diffusion, NSD vs. time for random movement is proportional to diffusion. Therefore, results are presented in the form of the estimated rate of diffusion, in $km^2\ day^{-1}$, which is calculated by dividing the slope of NSD vs. time by 4 for movement in 2 dimensions (Börger & Fryxell 2012).

For the site fidelity dispersal pattern, we quantified timing, duration, and distance traveled during forays that occurred during the summer with a non-linear mixed-effects model:

$$NSD_{ij} = \frac{\delta + d_i}{1+\exp\left(\frac{\theta_m - t}{\phi + f_i}\right)} + \frac{-\delta + d_i}{1+\exp\left(\frac{(\theta_r + r_i) - t}{\phi + f_i}\right)} + \varepsilon_{ij} \quad \text{(eq 1.5)}$$

$$d_i \sim N(0,\sigma^2_d),\ f_i \sim N(0,\sigma^2_f),\ r_i \sim N(0,\sigma^2_r),\ \varepsilon_{ij} \sim N(0, \sigma^2)$$

where δ, θ_m, θ_r, and φ are fixed-effects parameters for the asymptote (e.g., migration distance), date of migration, date of return, and scale (see Table 1.2) and d_i, f_i, and r_i are random-effects variables assumed to be normally distributed with mean 0 that represent individual variation in the asymptote, scale, and date of return. Within-group error ε_{ij} is assumed to be independent and normally distributed. The estimate for distance traveled during the foray is reported as the square root of the asymptote, δ. As approximately 95% of the distance between the midpoint of the migration and arrival at the new location occurs over the time span of $3^* \varphi$ (Börger & Fryxell 2012), timing of migration is estimated by $\theta_m - (3^* \varphi)$ and timing of return by $\theta_r + (3^* \varphi)$. Population estimates of average foray duration are calculated as the difference between the two: $(\theta_r + (3^* \varphi)) - (\theta_m - (3^* \varphi))$.

Selection of random-effects variables for the site fidelity model was conducted by first examining the range of coefficient values for each parameter, based on separate fits of the model to each trajectory, and selecting parameters with large variation as random effects in the full model (Pinheiro & Bates 2000). Alternative models with fewer random-effects variables and autocorrelation structures were tested against the null model using maximum likelihood and compared using AIC and likelihood ratio tests. The best model (eq. 1.5) also included an AR1 autocorrelation coefficient of 0.3. All mixed-effects models assumed a Gaussian error structure and were fit using the library "nlme" in the R program. The assumption of a normal distribution in random-effects estimates was checked using the Shapiro-Wilks test for normality.

1.3.6 Habitat relationships

We characterized habitat occupied by tagged halibut during the residential movement state using several habitat metrics available for the study area (depth, slope, habitat complexity, rugosity, substrate type, tidal velocity). Because more than 90% of the tagged animal observations occurred within a large area of the central portion of the bay that was characterized by a multibeam survey in 2001, we were able to use fine-scale depth (Carlson et al. 2002) and habitat information (Harney et al. 2006) resulting from this survey. However, observations for two tagged animals were removed from habitat analyses because a majority of their observations fell outside of the multibeam study area. Continuous rasters for slope,

change-in-slope (an indicator of slope interfaces and measure of habitat complexity), and rugosity (a measure of surface roughness) were derived from 5 m resolution depth data using ArcGIS 10.0 Spatial Analyst and ArcGIS 10.1 Benthic Terrain Modeler (Wright et al. 2005, ESRI 2011). Continuous rasters for soft sediment and moderate habitat complexity were derived from discrete habitat map polygons by calculating Euclidean distance from each grid cell to each type of polygon. Continuous information on time and depth averaged (monthly) tidal velocity was available from a 2-dimensional circulation model (ADCIRC) of Glacier Bay (Etherington et al. 2007b, Hill et al. 2009). We used information from these seven continuous rasters (depth, slope, change-in-slope, rugosity, distance from soft sediment, distance from moderate complexity habitat, and tidal velocity) to identify habitat associations and quantify the effects of habitat variables on home range size. Study area maps and additional details on habitat raster characteristics are available in Supplement1-3.

1.3.6.1 Habitat associations

To provide a simple description of the predominant habitat characteristics observed for the residential movement state relative to all available habitat types in the study area, we adapted an approach used to detect habitat associations based on the spatial distribution of catch during trawl surveys (Perry & Smith 1994). This method involves comparing the cumulative distribution function (CDF) of habitat values (e.g., depth) where tagged animals were observed to the CDF of available depths in the study area. Because halibut are large-bodied animals capable of a high degree of movement, we assumed they could have moved anywhere in the study area over the course of the observation period. To obtain CDFs for available habitat in the study area, a 20 m grid of the study area was created in ArcGIS (1.08E06 points) and values from each habitat raster were extracted at each grid point.

To account for telemetry error in the habitat analyses, a buffer with a radius of the estimated error was drawn around each tagged fish observation and all grid values within the buffer were averaged. "Observed" CDFs were then calculated using the median value of all observations in each home range to avoid pseudoreplication from treating repeated, irregular observations of one animal at one location as independent events (Rogers & White 2007). Confidence intervals for observed CDFs were generated by

bootstrapping, where the median observation for each home range was sampled with replacement 1000 times, and the 0.975 and the 0.025 quantile values were selected as the upper and lower confidence intervals.

We defined habitat associations by quantitatively comparing the CDFs for observed and available fish habitat. Specifically, for each habitat variable, we used the bootstrapped confidence levels for the observed CDFs to test for differences between observed and available habitat using the Kolmogorov-Smirnov (K-S) test. The K-S test is frequently used to test for differences between CDFs based on the maximum vertical difference (D) between the CDFs (Conover 1999). To determine whether positive differences existed between the observed and available CDFs (e.g., an association with shallower depths), we found the greatest positive difference (D^+) between the upper C.I. of the observed CDF and the available CDF. To determine whether negative differences existed between the observed and available CDFs (e.g., an association with deeper depths), we found the greatest negative difference (D^-) between the lower C.I. of the observed CDF and the available CDF. We determined D^+, D^-, and p values for each habitat variable using one-tailed Kolmogorov-Smirnov tests.

1.3.6.2 Habitat and home range size

We used a Generalized Additive Model (GAM) to determine whether home range sizes were related to habitat variables or fish total length. Intercept coefficients from the mixed-effects model for the residential movement state (in log format) were used as the response variable. Explanatory habitat variables were selected from the seven continuous habitat rasters used for habitat association analyses. In addition to habitat variables, we also included fish total length and year of study as explanatory variables for home range size. The GAM approach was used to allow for potential non-linearities in the relationship between response and explanatory variables. Prior to analysis, all variables were checked for covariance with the Pearson correlation coefficient; if a set of variables were found to be correlated, only the variable with the strongest relationship with the response variable was used in the model. After assessing correlation and linearity of habitat variables, two full models were tested:

Model 1: $y = \alpha + s_1(\text{depth}, k = 2) + s_2(\text{fish total length}, k = 3) + s_3(\text{distance from moderate complexity}, k = 3) + \beta_1 \text{change-in-slope} + \beta_2 \text{tidal velocity} + \text{year} + \varepsilon$ (eq 1.6)

Model 2: $y = \alpha + s_1(\text{depth}, \text{fish total length}, k = 3) + s_3(\text{distance from moderate complexity}, k = 3) + \beta_1 \text{change-in-slope} + \beta_2 \text{tidal velocity} + \text{year} + \varepsilon$ (eq 1.7)

where y is the vector of estimated intercepts from the mixed-effects model for all residential trajectory segments (n = 29), α and β_i are regression coefficients, s_i are smooth functions of the predictor variables, and ε are the residuals, assumed to be independent and normally distributed. GAM models were fit using maximum likelihood methods with a Gaussian error structure in the mgcv package in R (Wood 2006). Variables were sequentially removed from a full model based on the highest p value (i.e., larger than 0.05) and the best model was chosen based on the AICc criterion and residual analysis.

1.4 Results

1.4.1 Fish tagging and tracking

A total of 43 fish were tagged between 1991 and 1993 (Table 1.1). Most fish were tagged between June and September of each year, but 4 fish were tagged in November 1992 and tracked during the following summer. Tagged fish averaged 133 ± 32 cm SD in length and were released in good condition. Almost all (16 of 18) of the fish that we were able to sex were female, however sex could not be determined for the majority (n = 25) of tagged halibut in this study. Based on size and maturity ogives from International Pacific Halibut Commission (IPHC) records during this time period, the majority of the fish tagged in this study were likely to be adult females (Tim Loher, IPHC, pers. comm.).

Five fish were never relocated following tagging (Table 1.1). For the remaining 38 animals, the number of relocations ranged from 1 to 49 (average 17.4 ± 14.3 SD) per animal. More than half of the relocations for individual animals were obtained within 3 days of the previous observation, and 90% of the subsequent observations in each tracking period were within 8 days of the previous observation. Thus, the temporal scale of tagged animal observations during each tracking period can be characterized as

daily to weekly. In total, 706 acoustic tracking position estimates were obtained for all tagged animals in all years. Tracking effort differed among years, with most intense tracking during 1991 (32.5 observations per fish) and decreasing during 1992 (13.4 observations per fish) and 1993 (14.8 observations per fish). Tagged fish were observed over a mean tracking duration of 79.5 days (range 1 – 290 days) each year.

The average distance that individual tagged animals (n = 38) moved between the release location and the location of the last observation was 3.5 ± 0.8 km SE. The average maximum distance traveled during the entire observation was 5.8 ± 0.9 km SE. The maximum distance from release location recorded during the study was 17.9 km (Table 1.1). There were no significant relationships between the maximum distance traveled for each fish and fish size (linear regression, $p = 0.709$), tag size: body weight ratio (linear regression, $p = 0.637$), tag size (small vs. large; ANOVA, $p = 0.146$), or tag attachment method (interval vs. external; ANOVA, $p = 0.797$).

1.4.2 Movement states

The residential movement state was observed most frequently (27 of 43 tagged halibut). A total of 31 residential movement sequences (some animals had more than 1 residential sequence) were observed with a median duration of 58 days. The mixed-effects model population estimate for the intercept of NSD vs. time was 12.0 ± 0.3 m^2 SE, with a standard deviation for random effects of 1.4 on the log scale (Figure 1.4A). This corresponds to an estimated population home range radius of 401.3 m (95% C.I. 312.2 – 515.9 m) and 95% C.I.s for individual home range radii that range from 104.3 m to 1493.9 m on the untransformed scale.

The dispersive movement state was observed for 15 of 43 tagged halibut (Table 1.1, Fig. 1.3B). A total of 18 dispersive movement sequences were observed with a median duration of 27 days. This duration was significantly shorter than that of the residential movement state (t-test, $p < 0.0001$). The average maximum distance from the release location for fish that exhibited the dispersive movement state was 10.9 km.

Based on a randomization test with 1000 permutations, the median daily movement step length for observations from the residential movement state (330.4 m, n = 193 observations, Figure 1.5A) was

significantly less ($p < 0.0001$) than the median daily movement step length for observations from the dispersive movement state (861 m, n = 19 observations, Figure 1.5B). The rate parameter for the exponential curve that was fit to each step length distribution for use in the CRW analyses was 0.000213 ± .000015 SE for observations from the residential movement state and 0.0008059 ± 0.000184 SE for observations from the dispersive movement state.

The CRW simulations highlighted major differences in the scale and nature of dispersal between the residential and dispersive movement patterns. Average values of NSD observed for the dispersive movement state were generally within the 95% C.I.s for the CRWs, however these tended to be closer to the upper confidence level for the first 20 days (Figure 1.6A). Like the CRWs, observed values of NSD for the dispersive movement state exhibited a general trend for increased NSD values over time. In contrast, observed values of NSD vs. time for residential movement pattern were located along the lower 95% confidence interval for the simulations, and some observed values were smaller than the CRW confidence intervals after about 30 days (Figure 1.6B).

1.4.3 Individual dispersal patterns

Of the 38 fish that were relocated at least once following release, 32 produced a sufficient number of observations to allow classification of their dispersal pattern. More than half of these animals (n = 17) remained in the vicinity of the release location and demonstrated a home range (HR) dispersal pattern, where the average net displacement over average tracking durations of more than 3 months was less than 1 km (Table 1.3). Five animals demonstrated shifted home range dispersal patterns (SHR, n = 3 and FSHR, n = 2) with average net displacements of 10 and 4 km, respectively, over similar time periods. Six animals demonstrated the site fidelity (SF) dispersal pattern by moving an average maximum displacement of approximately 10 km but eventually returning to locations that were an average net displacement of less than 2 km from the release location over average time periods of more than 4 months. Two of these fish established home ranges at other locations in the study area during the foray. Three fish classified with the site fidelity dispersal pattern returned to within hundreds of meters (range 200 – 500 m) of their release locations after moving an average maximum distance of 9.8 km ± 0.5 km SE. One animal

classified as SHR (ID #5, Figure 1.2) also demonstrated a site fidelity pattern during a temporary foray of 6 km and duration of 16 days followed by a return to within 200 m of the location occupied prior to the foray. Four fish exhibited the random (R) dispersal pattern, moving an average maximum and average net displacement of approximately 12 km over observation periods that averaged less than 2 months (approximately half of the typical durations observed for the animals assigned to other dispersal patterns). The six fish for which only a few observations were collected (U) had very short observation durations (average 13.5 days), yet the average net displacement of approximately 3 km for these fish was greater than that observed for the home range dispersal pattern. There was no significant difference among the total lengths of fish in each of the five dispersal patterns, unclassified movements, or the animals that were never observed after tagging (Kruskal-Wallis test, $p = 0.3679$, $df = 6$).

Two of the six tagged fish that were classified as having a site fidelity dispersal pattern were not included in the mixed-effects model, as their movements occurred over the winter. Both of these animals were tagged in November and tracked again the following summer, so the timing of their movements could not be compared to animals that were observed only during the summer. For the four remaining animals in this group, the population estimate for the distance traveled during their forays (the asymptote, δ) was a movement radius of 10.5 km ± 4.0 km SE ($p < 0.0001$) (Figure 1.4B). The standard deviation for individual animals from the population mean (d_i) was 5.1 km. The scale parameter φ was estimated to be 2.4 ± 0.5 days SE ($p < 0.0001$), and the standard deviation for the corresponding random-effects variable f_i was 0.0001 days. The population estimate for the timing of departure was July 8 ± 1 day SE ($p < 0.0001$), and the estimate for the timing of the return to the original location was August 17 ± 5 days SE ($p < 0.0001$) with a standard deviation for the random-effects variable r_i of 10.1 days. The overall population estimate for duration of the forays was 40 ± 5.1 days SE. The value of the AR(1) autocorrelation coefficient (referred to in the nlme package as Phi) for the model was 0.61.

The population (fixed effect) estimate of the slope of NSD vs. time for the random dispersal pattern was 3.6 ± 0.2 km day^{-1} SE ($p < 0.0001$) which corresponds to an estimated diffusion constant D of 0.9 ± 0.05 km^2 day^{-1} SE (Figure 1.4C). In contrast to the home range dispersal pattern, there was very

little difference between the individual estimated movement rates because most of the variation was attributed to residual variation around the fixed effect.

1.4.4 Habitat relationships

1.4.4.1 Habitat associations

Significant habitat associations were observed between tagged animals in the residential movement state and available habitat in the study area (Table 1.4, Figure 1.7). Relationships were strongest for the habitat heterogeneity variables of change-in-slope, slope, and rugosity, with tagged animals tending to occupy areas of higher habitat heterogeneity relative to the range of values available in the study area. Home ranges were also associated with intermediate values of tidal velocity, where significant differences between observed and available habitat occurred at both high and low values of tidal velocities. Finally, home ranges were associated with shallower depths relative to the range of available depths in the study area, with approximately 75% of home ranges occurring in depths less than 100 m. No significant differences were observed between tagged fish and distance to moderate complexity or distance to soft substrate variables.

1.4.4.2 Habitat and home range size

The results from the GAM analysis indicate that of the variables examined, home range size varied most strongly with depth. The best model contained only a depth term with an estimated degrees of freedom of 1.8 ($p = 0.007$). Home range size increased with increasing depth for depths < approximately 150 m (Figure 1.8). The deviance explained by the selected model was 30.0 %. No other models with all significant terms were within +/- 2 ΔAICc of the best model. No patterns were observed in the residuals, which were not non-normally distributed (Shapiro-Wilk test, $p = 0.65$).

1.5 Discussion

Although halibut are large-bodied fish capable of moving thousands of kilometers during winter spawning migrations (Skud 1977, Loher & Seitz 2006), our results suggests that limited dispersion at very small spatial scales may be a common phenomenon for adult female halibut in Glacier Bay during the summer and into the fall. The residential movement state was demonstrated by the majority (27/43) of the

fish tagged in this study. The home range dispersal pattern (which consists of residential movement in the vicinity of the release location throughout the observation period) was also the most frequently observed dispersal pattern (n = 17) among the 32 individual animals for which dispersal patterns could be determined. Although animals that exhibited the dispersive movement state moved more broadly around the study area, these movements were still relatively small (< 20 km) compared to the distances moved during winter migrations.

Animals that were never relocated or were relocated too infrequently to characterize their movement patterns (11/43 fish) may have exhibited a more mobile movement pattern and thus moved out of the study area quickly. In this case, they could have moved to areas within Glacier Bay that were not monitored during acoustic surveys or they could have left the inside waters of Glacier Bay entirely. Alternatively, they may have been captured in commercial harvests that were occurring in Glacier Bay, experienced mortality, or the tag could have been shed or ceased to function.

1.5.1 Movement states

Telemetry records often document different behaviors among individuals that are driven by different movement "states" (Blackwell 1997, Morales et al. 2004). For example, a period of intensive foraging may result in a movement state with little net displacement, while a period of migration may result in a movement state with relatively large net displacement. Typically, ecologists are interested in the spatial and temporal scales of these movement states, as well as habitat attributes with which they may be associated (Papworth et al. 2012).

The two movement states, residential and dispersive, that tagged fish exhibited during the summer and fall differed in terms of scale, duration, and potential for dispersion. The residential movement state was associated with average movement scales of less than 1 km for several months at a time and a sustained, non-random lack of dispersion. In contrast, the dispersive movement state was characterized by greater spatial scales (approximately 10 km), shorter temporal durations (< 1 month), and likely contained a mix of random and directed movement.

Because tracking occurred during the summer foraging season, and large adult halibut have few predators, these two movement states could reflect different underlying foraging strategies. Both "sit-and-wait" ambush and active searching are common foraging tactics for flatfish species (Gibson 2005). The residential movement pattern could be driven by a sit-and-wait tactic, which would require little movement in areas where prey is delivered to the fish. Based on laboratory studies, a closely related congener Atlantic halibut *Hippoglossus hippoglossus* is thought to be an ambush predator that employs a sit-and-wait feeding tactic (Haaker 1975, Nilsson et al. 2010). Other flatfish such as summer flounder *Paralichthys dentatus* have been observed to employ a variety of foraging tactics, including ambush and active pursuit, that change with prey type (Staudinger & Juanes 2010). Thus, switching between the two movement states may occur in conjunction with changes in the type, abundance, distribution, and mobility of prey species (see Nakano et al. 1999). However, the dispersive movement pattern could also include animals that are moving in a directed manner from one feeding location to another.

1.5.1.1 Caveats

It is important to emphasize that the data presented in this study are inherently positive and biased toward the observation of the residential movement state. The experimental design employed in this study, which featured searching for tagged animals in the vicinity of their last known location, resulted in a much better characterization of the residential movement state compared to the dispersive movement state. The tracking procedure was effective for locating tagged animals that were occupying home ranges, as the detection range for the acoustic tags (up to 2 km) was larger than the scale of most home ranges. However due to the difficulty of tracking more mobile animals for long time periods in the large study area, it is likely that the dispersive movement state occurred more frequently than was observed and its spatial extent was not fully characterized. Although detection ability was adequate for characterizing the residential state throughout the study, changes in tag size and frequency (Supplement 1-1) could have improved the detection of animals in the dispersive movement state as the study progressed. It is likely that the largest movement observed (18 km) probably reflects a practical limit for the area searched during this study, so movements beyond that would have a low probability of detection.

Although it is possible that unknown tagging effects may have affected the behavior of tagged halibut, we feel that tagging effects are unlikely to have affected the scale and nature of halibut movement reported in this study for several reasons. First, a long-term laboratory study of both internal and external archival tag attachment suggests that both types of attachment are well-tolerated by halibut and have not resulted in changes in behavior compared to controls (Loher & Rensmeyer 2011). Second, the tags were small relative to the size of the fish (average = 0.1%, maximum = 0.4%). Third, Pacific halibut fitted with much larger pop-up satellite tags have been observed to move more than 1000 km (Loher & Seitz 2006). Finally, we found no statistical relationships between fish size or tag:body size ratio and maximum displacement and no relationship between maximum displacement and tag size (small or large) or type of attachment (internal or external).

1.5.2 Individual dispersal patterns

1.5.2.1 Non-random dispersal patterns: home range and site fidelity

The majority of tagged fish in this study exhibited distinctly non-random individual dispersal patterns that were dominated by home range but also included temporary long-distance forays followed by return to previously occupied locations and shifting of home ranges to new locations. The prominence of the home range dispersal pattern suggests that regular use of relatively small areas could be a common phenomenon during summer. Several acoustic telemetry studies have demonstrated summer home range behavior for other flatfish species such as adult English sole *Parophrys vetulus* (Sandra O'Neill, Washington Dept. of Fish and Wildlife, pers. comm), and juvenile California halibut *Paralichthys californicus* (Espasandin 2012) that occurs at scales < 1 km. Because some halibut shifted locations for home range behavior, it is possible that some fish may switch home range locations depending on changes in prey distribution and abundance and thus may not have fidelity to specific locations.

However, multiple observations of tagged animals returning to within several hundred meters of previously occupied locations following larger-scale movements (e.g., 10 km distance, 1 month duration) suggests that some halibut do have site fidelity to specific locations (as defined by Giuggioli & Bartumeus 2011). The site fidelity dispersal pattern was observed for 7/43 animals (including one animal,

ID #5, that was assigned to the shifted home range dispersal pattern). It is also possible that temporary departures from home ranges were not detected due to the irregular nature of the tracking trips and the difficulty of relocating wide-ranging fish. In that case, subsequent relocation of these same individuals at previously occupied locations would indicate intra-annual site fidelity to established home ranges. Therefore intra-annual site fidelity may be a key feature of adult female halibut movement patterns in Glacier Bay during the summer and fall.

The study has also provided some evidence for inter-annual site fidelity for halibut in Glacier Bay. Of the four animals released in November, three inhabited home ranges at their release locations the following summer. Whether or not these animals left Glacier Bay during winter spawning migrations is unknown, but two of these animals were observed at different locations within the park following tagging (thus demonstrating a site fidelity dispersal pattern).

These results complement previous observations of site fidelity for Pacific halibut from a pop-up satellite archival tag (PSAT) study and provide further details on the scales at which it may occur. Approximately 80% of summer-to-summer PSAT pop-up locations (n = 25) were located within 20 km of release locations after one year at liberty (Loher 2008). Most (75%) of these fish had returned to the release location following migrations to deeper water in the Gulf of Alaska during winter, presumably to spawn. Although the displacement from the release locations from the PSAT study matches the scale of the dispersive movement state observed in this study, the demonstrated ability of fish in the current study to return to within a few hundred meters of their original locations after undertaking forays indicates that site fidelity for Pacific halibut likely occurs at much finer spatial scales than can be detected using PSATs. Site fidelity has also been observed for many other flatfish species (Hunter et al. 2003, Solmundsson et al. 2005, Sackett et al. 2008, Dando 2011, Moser et al. 2013).

1.5.2.2 Random movement: diffusion

Although the majority of the fish in this study displayed non-random movement patterns associated with an overall lack of dispersal during summer, some fish did appear to have more mobile movement patterns. The random movement dispersal pattern demonstrated by a small proportion of

tagged halibut suggests that some halibut do not establish home ranges, but may instead move randomly throughout summer foraging areas. The rate of diffusion associated with random movement in this study, 0.9 km^2 day^{-1}, was comparable to diffusion rates estimated for several other flatfish species such as Baltic Sea turbot, *Psetta maxima* (Florin & Franzén 2010) and winter flounder, *Pseudopleuronectes americanus* (Saila 1961) based on results derived from conventional tag recaptures. Therefore, random movement appears to be another common behavior of Pleuronectiformes species, likely as a foraging tactic. However, sample sizes were low for this dispersal pattern, so results should be interpreted with caution. For example, these fish could also have been detected during temporary forays to or from home ranges in unknown locations.

A large-scale summer-to-summer PIT tag study of 67 000 halibut provided similar observations of both sedentary and mobile movement patterns for adult halibut that occurred over larger scales in space and time. Fish tagged during 2003 – 2004 had not mixed completely with the population by 2006 – 2009 (Webster et al. 2013) and as of 2008, 86% of 132 tags recaptured by annual survey vessels were caught at the same survey station where they were released (Loher 2008). Survey stations were located on an 18.5 km grid, which matches the approximate scale of the dispersive movement state observed in Glacier Bay. These observations support the presence of a long-term sedentary movement pattern for adult fish. On the other hand, the probability for large-scale movement between management units for large (e.g. 130 cm) fish was close to 20% for some units (Webster et al. 2013), which suggests that a more mobile movement pattern with a greater potential for dispersal also exists for adult halibut. In addition to the 4 fish that exhibited the random dispersal pattern in our study, it is possible that some of the 11 fish that were rarely or never detected had more mobile movement patterns. In that case, the proportion of tagged animals with more mobile patterns would range from a minimum of 9% (4/43) to a maximum of 35% (15/43), assuming no mortality, tag loss, or undetected home range behavior at unknown locations within Glacier Bay had occurred.

1.5.2.3 Caveats

Our use of a model selection framework to link observed patterns of NSD vs. time to theoretical models of dispersal represents a promising approach for identifying and quantifying animal movement patterns in terms of ecological phenomena such as home range occupation, foraging, and migration. This analysis method is appropriate for data collected at irregular intervals because the analysis is based on positive observations of NSD at a given point in time. However, due to the small sample sizes obtained for animals with more mobile movement patterns in this study, the results for dispersal patterns other than home range should be viewed as providing a preliminary understanding of the types of behavior and spatial scales of movement that fish may demonstrate during summer.

1.5.3 Habitat relationships

The habitat associations observed for tagged fish may be related to a tendency for tagged fish to occupy a specific benthic habitat type in Glacier Bay. Three regions composed of different combinations of depth, tidal velocity, substrate type, and community composition exist in Glacier Bay (Etherington et al. 2007a). The mouth and lower portions of Glacier Bay consist of a large, flat, shallow (50 m), high-current area with sand and cobble substrate associated with a community of horse mussels, scallops, and sea urchins. In contrast, the central and northern portions of the bay are composed primarily of deep fjords (to approximately 450 m) with muddy substrates (Supplement 1-3, Figure S1.3-1) and were associated with Tanner crab (*Chionoecetes bairdi*), shrimp, and flatfish species. However, the majority of animals in this study were tagged and tracked in a transition zone between these two areas that is characterized by intermediate depths, intermediate levels of tidal velocity, mixed cobble/soft sediment, and intermediate to high levels of habitat complexity. This region is also occupied by Pacific herring (*Clupea palasii*), Pacific cod (*Gadus macrocephalus*), walleye pollock (*Gadus chalcogrammus*), rockfishes (*Sebastes* spp.), and other common prey items for halibut (Best & St-Pierre 1986, Etherington et al. 2007a, Moukhametov et al. 2008, Renner et al. 2012). This transition area is also a highly productive front where well-mixed water from the mouth of the bay meets nutrient-rich stratified waters from the fjords (Etherington et al. 2007b).

Significant associations between the residential movement state and measures of habitat heterogeneity (change-in-slope, rugosity) and tidal velocity may also be related to a sit-and-wait foraging strategy. For example, complex habitat can aid concealment during ambush and tides may deliver pelagic prey (see Beaudreau & Essington 2011) to ambush predators. The strongest habitat association observed was for the change-in-slope variable, which represents interfaces between shallow and steep slopes as well as areas where depth is frequently changing. Associations with interfaces between different habitat types have been observed for other fish species such as the barred sand bass, *Paralabrax nebulifer*, which inhabited interfaces between rocky reefs used for hunting and adjacent soft-sediment habitats used for resting or refuge (Mason & Lowe 2010). Although flatfishes are often associated with soft sediments related to their tendency to bury in sediments (Gibson 2005), no significant habitat association with distance to soft substrate habitats was observed here, a result that could be related to the abundance of soft sediment in the study area or a reduced tendency to bury in sediment for adult fish compared to juveniles.

Of the environmental and biological explanatory variables examined, only depth was significantly related to home range size. Increased scales of movement in deeper areas could reflect differences in prey type or prey densities compared to shallow areas (i.e., transition region communities compared to deep fjord bottom communities, as discussed above). A positive relationship with depth has also been observed for temperate reef-associated fishes, where species that occupy deeper depths tend to have larger ranges of movement than those that occur at shallower depths (Freiwald 2012). An increase in telemetry error with depth could confound the relationship between depth and home range size, however the expected telemetry error is still small (100 m) relative to the distances moved in the larger home ranges (1 – 2 km). Note that the 95% C.I.s reported for the GAM do not include errors associated with uncertainty of position related to depth. Because this model explained only 30% of the variance, home range is probably affected by variables that were either not measured or occurred at different scales in space or time. For example, the movements of many fish species are known to be related to tidal patterns on a daily basis (Tolimieri et al. 2009), so the use of time-and-depth averaged tidal velocity in this study may have been too coarse to detect relationships with tide. Although positive relationships between fish length and home

range size have been previously reported (Kramer & Chapman 1999), we found that home range size was not related to fish size in this study. However the size distribution of the fish tagged in this study was relatively homogeneous (Supplement 1-1, Figure S1.1-1), so this result could also be related to low numbers of very small or very large fish. Finally, home range size did not change over time based on the lack of significance of the year variable. This result implies a stability of home range scales over time as well as a lack of effect of changing tag types (frequency, size, longevity, attachment method) as the study evolved (Supplement 1-1, Table S1.1-2). Recent fieldwork by the authors (Chapter 2), where 15 adult female halibut were tracked for 2 months in Glacier Bay, has provided independent confirmation of the scale and dominance of the residential movement pattern during summer.

1.5.4 Implications for MPAs and spatial fisheries management

Determining fish movement scales relative to MPA size is one of the most important aspects of MPA design (Gruss et al. 2011, Saarman et al. 2013). Scales of both residential and dispersive movement states were smaller than the scale of the Glacier Bay MPA, and most tagged animals were detected regularly inside the MPA boundary[2]. Thus, Glacier Bay is likely to encompass the majority of movements of individual adult female Pacific halibut during the summer and fall. Retention of halibut within Glacier Bay may be encouraged by the enclosed nature of the bay, availability of heterogeneous habitat with which tagged halibut were associated, and the productivity of its waters. Glacier Bay would therefore be expected to serve as a refuge from commercial harvest after the phase-out of commercial fisheries in the park is completed (estimated by the National Park Service to occur sometime between 2040 and 2050). However, understanding the potential for specific benefits of Glacier Bay as an MPA such as change in size structure or abundance (Taggart et al. 2004) will require more information on 1) large-scale movement patterns of halibut over yearly timescales, as fish could be vulnerable to commercial fishing during migrations from summer foraging locations to winter spawning locations outside of Glacier Bay,

[2] The technical boundary for the Glacier Bay National Park Marine Protected Area extends to the outside waters, but in this document we refer to the functional MPA of the interior waters, known as "Glacier Bay Proper", within which commercial fishing and vessel traffic are regulated by the National Park Service.

and 2) the effect of on-going charter and unguided sport fishing within the park that will continue after commercial fishing is phased out.

In addition to insights into the potential utility of Glacier Bay as an MPA, this study has also yielded information that may be useful for design of MPAs or spatial management in general for halibut. The frequent observations of non-random dispersal patterns such as home range and site fidelity are strikingly similar to dispersal patterns observed for other temperate reef-associated fishes such as lingcod *Ophiodon elongatus* and some rockfish species (Matthews 1990, 1992, Pearcy 1992, Starr et al. 2004, Tolimieri et al. 2009, Beaudreau & Essington 2011). Lingcod, yellowtail rockfish *Sebastes flavidus*, blue rockfish *Sebastes mystinus*, and California scorpionfish *Scorpaena guttata* were found to exhibit non-dispersive movement at spatial scales that were similar to tagged halibut in this study (Freiwald 2012, Supplement 1, Figure S1). Benefits such as increases in biomass and egg production have been observed for multiple temperate reef fish species within MPAs in California (Tetreault & Ambrose 2007). Movement patterns and scales for adult female lingcod tagged in a small MPA near Sitka, Alaska, which were very similar to those of the adult female halibut tagged in our study, suggest that MPAs may facilitate increased egg production through protection of lingcod brood stock (Starr et al. 2004). Our research suggests that MPAs may also provide some degree of protection for Pacific halibut brood stock, and therefore potential enhancement of egg production, based on observations of home range and site fidelity during the summer and fall. The effectiveness of such MPAs would then depend on the timing of winter spawning migrations, which may occur on the shoulders of the commercial fishing season (Loher 2011), as well as the proportion of adult females that undertake annual spawning migrations (Loher & Seitz 2008). As the estimated total biomass of Pacific halibut in the eastern North Pacific Ocean has declined by 50% between 1996 and 2013 and the majority of fish captured in the commercial fishery are females (Stewart et al. 2013), additional time-area closures in areas where large females are found in high abundance could provide some measure of protection during periods of declining stock abundance.

However, additional research will be required to determine the extent to which the movement scales and habitat associations observed in Glacier Bay, a relatively enclosed fjord estuary, may be

applicable to other marine environments that occur throughout the range of halibut distribution. For example, limited movement within fjords has been observed for Atlantic halibut (Seitz et al. 2014) and Atlantic cod (Hedger et al. 2011). Movement scales may also be related to the availability of complex habitat and high relief areas in a given location (Matthews 1990, Beaudreau & Essington 2011). Our results indicate that in Glacier Bay, scales of movement tend to be smaller in shallow (< 100 m) areas with heterogeneous topography and complex habitat compared to deep, flat areas with low levels of habitat complexity. Thus, although some locations are likely to be more effective than others at retaining the movements of halibut per unit area, more research on halibut movement scales in different habitat types and geographic regions is needed to determine how much of the sedentary movement patterns observed in this study are due to 1) fjord topography, 2) the presence of complex habitat, or 3) inherent behavior of adult halibut.

Finally, the existence of fine-scale site fidelity for some proportion of reproductive females has implications for depletion of mature female fish at local scales. Although the phenomenon of local depletion for Pacific halibut has not been explicitly documented, some evidence from a variety of sources suggests that it may be of concern. For example, declines in commercial catch per unit effort near the Pribilof Islands occurred in conjunction with concentrated local fishing effort (Hare 2005). Localized declines have also been observed near populated areas, where intense charter and sport fishing effort occurs (Trumble et al. 1991). Our findings of home range behavior combined with site fidelity and the ability to return to previously occupied areas provide a mechanism by which Pacific halibut could be vulnerable to local depletion. However, before the potential for local depletion can be characterized, broad geographic-scale dispersal processes, population connectivity, and spatial structure during other life history phases, such as passive planktonic larval drift and contranatant ontogenetic migrations, must be fully assessed (Skud 1977, Conners & Munro 2008).

1.5.5 Benefits of NSD analysis methods

Our use of NSD analyses to describe and quantify movement patterns and scales is a novel, robust approach that can be applied to irregular datasets commonly collected for fishes. Using the analysis of

dispersal patterns, telemetry records for tagged animals can be analyzed in the context of movement ecology and potential for dispersal rather than focusing on the size or location of an area inhabited by the animal. The mixed-effects model framework allows the description of inherent individual variability as well as borrows strength from individuals with more observations. Information on movement characteristics is provided in formats such as diffusion rates or home range scales that can be easily used to simulate movement paths for use in other studies. For example, such simulations could be used to estimate energy budgets during residential versus dispersive movement states. In addition, it provides a way to compare results between acoustic and conventional tagging studies. For example, diffusion coefficients have been calculated for American lobsters using the slope of NSD vs. time from conventional tag recovery data (den Heyer et al. 2009). Thus, this method may provide a way to leverage the detailed information provided by acoustic studies with the larger sample sizes available from conventional tagging studies

Finally, the NSD analysis framework complements information needs for MPA design, where managers are often faced with the task of compiling information on the movement scales of multiple fish species based on data collected using different methods (Saarman et al. 2013). Because home range scales are reported as a movement radius rather than area, NSD methods can be used to compare scales for fish that move in one dimension (e.g., in a river or along a coastline) with those that move in two dimensions. Combining different types of movement data (e.g., large-scale acoustic arrays, archival tags, or conventional tags) would also be possible because the only information required for this method is distance moved from the release location over time. Therefore this analysis approach may be particularly valuable because it provides movement data in formats that can be easily combined or compared with results from other studies.

1.6 Acknowledgements

We would like to thank the National Park Service, the U.S. Geological Survey, the Rasmuson Fisheries Research Center, the Pollock Conservation Cooperative Research Center, and the North Pacific Research Board for financial support of this project. Many people assisted with the fieldwork and data

processing for this project, including Jim de La Bruere, Elizabeth Hooge, Gretchen Bishop, Liz Chilton, Caroline Dezan, Fritz Koschmann, Liz Solomon, and others. We thank Milo Adkison, Franz Mueter, Tim Loher, Susanne McDermott, Anne Beaudreau, Chad Soiseth, and Craig Murdoch for providing valuable advice and guidance on the manuscript.

1.7 Literature cited

Austin D, Bowen WD, McMillan JI (2004) Intraspecific variation in movement patterns: modeling individual behaviour in a large marine predator. Oikos 105:15-30

Bartumeus F, Da Luz MGE, Viswanathan GM, Catalan J (2005) Animal search strategies: a quantitative random-walk analysis. Ecology 86:3078-3087

Beaudreau AH, Essington TE (2011) Use of pelagic prey subsidies by demersal predators in rocky reefs: insight from movement patterns of lingcod. Mar Biol 158:471-483

Best EA, St-Pierre G (1986) Pacific halibut as predator and prey. Int Pac Halibut Comm Tech Rep 21:1-27

Blackwell PG (1997) Random diffusion models for animal movement. Ecol Model 100:87-102

Börger L, Fryxell J (2012) Quantifying individual differences in dispersal using net squared displacement. In: Clobert J, Baguette M, Benton T, Bullock J (eds) Dispersal and spatial evolutionary ecology. Oxford University Press, p 222 - 230

Bunnefeld N, Börger L, van Moorter B, Rolandsen CM, Dettki H, Solberg EJ, Ericsson G (2011) A model-driven approach to quantify migration patterns: individual, regional and yearly differences. J Anim Ecol 80:466-476

Burnham KP, Anderson DR (1990) Model selection and multimodal inference: an information-theoretic approach. Springer Verlag, New York

Carlson PR, Hooge P, Cochrane G, Stevenson A, Dartnell P, Lee K (2002) Multibeam bathymetry and selected perspective views of main part of Glacier Bay, Alaska. Report No. 02-391, U.S. Geological Survey, Menlo Park, CA

Clark WG, Hare SR (2006) Assessment and management of Pacific halibut: data, methods, and policy. Int Pac Halibut Comm Sci Rep 83:1 - 104

Conners ME, Munro P (2008) Effects of commercial fishing on local abundance of Pacific cod (*Gadus macrocephalus*) in the Bering Sea. Fish Bull 106:281-292

Conover WJ (1999) Practical nonparametric statistics. John Wiley & Sons, Inc.

Dando PR (2011) Site fidelity, homing and spawning migrations of flounder *Platichthys flesus* in the Tamar estuary, South West England. Mar Ecol Prog Ser 430:183-196

den Heyer CE, Chadwick EMP, Hutchings JA (2009) Diffusion of American lobster (*Homarus americanus*) in Northumberland Strait, Canada. Can J Fish Aquat Sci 66:659-671

Espasandin C (2012) Movement behavior of the California halibut in a restored southern California wetland. M.S. thesis, California State University, Long Beach, CA

ESRI (2011) ArcGIS Desktop, Release 10.1, Redlands, CA

Etherington LL, Cochrane GR, Harney J, Taggart SJ, Mondragon J, Andrews AG, Madison E, Chezar H, de La Bruere LJ (2007a) Glacier Bay seafloor habitat mapping and classification - first look at linkages with biological patterns. In: Piatt JF, Gende SM (eds) Fourth Glacier Bay Science Symposium. U. S. Geological Survey, Juneau, Alaska, p 71-75

Etherington LL, Hooge PN, Hooge ER, Hill DF (2007b) Oceanography of Glacier Bay, Alaska: implications for biological patterns in a glacial fjord estuary. Estuaries Coasts 30:927-944

Florin A-B, Franzén F (2010) Spawning site fidelity in Baltic Sea turbot (*Psetta maxima*). Fish Res 102:207-213

Freiwald J (2012) Movement of adult temperate reef fishes off the west coast of North America. Can J Fish Aquat Sci 69:1362-1374

Geiselman J, Dunlap J, Hooge P, Albert D (1997) Glacier Bay ecosystem GIS CD-ROM set. U.S. Geological Survey and Interrain Pacific, Anchorage and Juneau, AK.

Gibson RN (2005) The behaviour of flatfishes. In: Gibson RN (ed) Flatfishes: biology and exploitation. Blackwell Science Ltd, Oxford, UK, p 213-239

Giuggioli L, Bartumeus F (2011) Linking animal movement to site fidelity. J Math Biol DOI 10.1007/s00285-011-0431-7:1-10

Gruss A, Kaplan DM, Guenette S, Roberts CM, Botsford LW (2011) Consequences of adult and juvenile movement for marine protected areas. Biol Conserv 144:692-702

Haaker P (1975) The biology of the California halibut, *Paralichthys californicus* (Ayres) in Anaheim Bay. In: Lane ED, WHill CW (eds) The marine resources of Anaheim Bay, Vol 165. Calif. Dep. Fish Game Fish. Bull., p 137-151

Hare SR (2005) Investigation of the role of fishing in the Area 4C CPUE decline. In: Sadorus L (ed) Report of Assessment and Research Activities, 2004. International Pacific Halibut Commission, Seattle, WA, p 185-197

Harney JN, Cochrane GR, Etherington LL, Dartnell P, Golden NE, Chezar H (2006) Geologic characteristics of benthic habitats in Glacier Bay, southeast Alaska. Report No. 2006-1081, U.S. Geological Survey, Santa Cruz, CA

Hedger RD, Uglem I, Thorstad EB, Finstad B, Chittenden CM, Arechavala-Lopez P, Jensen AJ, Nilsen R, Okland F (2011) Behaviour of Atlantic cod, a marine fish predator, during Atlantic salmon post-smolt migration. ICES J Mar Sci 68:2152-2162

Hill DF, Ciavola SJ, Etherington L, Klaar MJ (2009) Estimation of freshwater runoff into Glacier Bay, Alaska and incorporation into a tidal circulation model. Estuar Coast Shelf Sci 82:95-107

Hunter E, Metcalfe JD, Reynolds JD (2003) Migration route and spawning area fidelity by North Sea plaice. P Roy Soc B:Biol Sci 270:2097-2103

Jorgensen SJ, Reeb CA, Chapple TK, Anderson S, Perle C, Van Sommeran SR, Fritz-Cope C, Brown AC, Klimley AP, Block BA (2010) Philopatry and migration of Pacific white sharks. P Roy Soc B:Biol Sci 277:679-688

Kareiva PM, Shigesada N (1983) Analyzing insect movement as a correlated random walk. Oecologia 56:234-238

Kramer DL, Chapman MR (1999) Implications of fish home range size and relocation for marine reserve function. Environ Biol Fishes 55:65-79

Loher T (2008) Homing and summer feeding site fidelity of Pacific halibut (*Hippoglossus stenolepis*) in the Gulf of Alaska, established using satellite-transmitting archival tags. Fish Res 92:63-69

Loher T (2011) Analysis of match–mismatch between commercial fishing periods and spawning ecology of Pacific halibut (*Hippoglossus stenolepis*), based on winter surveys and behavioural data from electronic archival tags. ICES J Mar Sci 68:2240-2251

Loher T, Rensmeyer R (2011) Physiological responses of Pacific halibut, *Hippoglossus stenolepis*, to intracoelomic implantation of electronic archival tags, with a review of tag implantation techniques employed in flatfishes. Rev Fish Biol Fish 21:97-115

Loher T, Seitz AC (2006) Seasonal migration and environmental conditions of Pacific halibut *Hippoglossus stenolepis*, elucidated from pop-up archival transmitting (PAT) tags. Mar Ecol Prog Ser 317:259-271

Loher T, Seitz AC (2008) Characterization of active spawning season and depth for eastern Pacific halibut (*Hippoglossus stenolepis*), and evidence of probable skipped spawning. J Northwest Atl Fish Sci 41:23-36

Mason TJ, Lowe CG (2010) Home range, habitat use, and site fidelity of barred sand bass within a southern California marine protected area. Fish Res 106:93-101

Matthews KR (1990) A telemetric study of the home ranges and homing routes of copper and quillback rockfishes on shallow rocky reefs. Can J Zool 68:2243-2250

Matthews KR (1992) A telemetric study of the home ranges and homing routes of lingcod *Ophiodon elongatus* on shallow rocky reefs off Vancouver Island, British Columbia. Fish Bull 90:784-790

Moorcroft PR, Lewis MA (2006) Mechanistic home range analysis. Princeton University Press, Princeton and Oxford

Morales JM, Haydon DT, Frair J, Holsinger KE, Fryxell JM (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology 85:2436-2445

Moser ML, Myers MS, E. West J, O'Neill SM, Burke BJ (2013) English sole spawning migration and evidence for feeding site fidelity in Puget Sound, U.S.A., with implications for contaminant exposure. Northwest Sci 87:317-325

Moukhametov IN, Orlov AM, Leaman BM (2008) Diet of Pacific halibut (*Hippoglossus stenolepis*) in the northwestern Pacific Ocean. Int Pac Halibut Comm Tech Rep 52

Nakano S, Fausch KD, Kitano S (1999) Flexible niche partitioning via a foraging mode shift: a proposed mechanism for coexistence in stream-dwelling charrs. J Anim Ecol 68:1079-1092

Nilsson J, Kristiansen TS, Fosseidengen JE, Stien LH, Ferno A, van den Bos R (2010) Learning and anticipatory behaviour in a "sit-and-wait" predator: the Atlantic halibut. Behav Processes 83:257-266

Nouvellet P, Bacon JP, Waxman D (2009) Fundamental insights into the random movement of animals from a single distance-related statistic. Am Nat 174:506-514

Papworth SK, Bunnefeld N, Slocombe K, Milner-Gulland EJ (2012) Movement ecology of human resource users: using net squared displacement, biased random bridges and resource utilization functions to quantify hunter and gatherer behaviour. Method Ecol Evol 3:584–594

Pearcy WG (1992) Movements of acoustically-tagged yellowtail rockfish *Sebastes flavidus* on Heceta bank, Oregon. Fish Bull 90:726-735

Perry RI, Smith SJ (1994) Identifying habitat associations of marine fishes using survey data: an application to the northwest Atlantic. Can J Fish Aquat Sci 51:589-602

Pinheiro J, Bates D (2000) Mixed-effects models in S and S-PLUS. Springer-Verlag, New York

Renner M, Arimitsu ML, Piatt JF (2012) Structure of marine predator and prey communities along environmental gradients in a glaciated fjord. Can J Fish Aquat Sci 69:2029-2045

Rogers KB, White GC (2007) Analysis of movement and habitat use from telemetry data. In: Brown M, Guy C (eds) Analysis and interpretation of freshwater fisheries data. American Fisheries Society, Bethesda, Maryland, p 625-676

Saarman E, Gleason M, Ugoretz J, Airamé S, Carr M, Fox E, Frimodig A, Mason T, Vasques J (2013) The role of science in supporting marine protected area network planning and design in California. Ocean Coast Manage 74:45-56

Sackett DK, Able KW, Grothues TM (2008) Habitat dynamics of summer flounder *Paralichthys dentatus* within a shallow USA estuary, based on multiple approaches using acoustic telemetry. Mar Ecol Prog Ser 364:199-212

Saila SB (1961) A study of winter flounder movements. Limnol Oceanogr 6:292-298

Seitz AC, Loher T, Norcross BL, Nielsen JL (2011) Dispersal and behavior of Pacific halibut *Hippoglossus stenolepis* in the Bering Sea and Aleutian Islands region. Aquat Biol 12:225-239

Seitz AC, Michalsen K, Nielsen JL, Evans MD (2014) Evidence of fjord spawning by southern Norwegian Atlantic halibut (*Hippoglossus hippoglossus*). ICES J Mar Sci doi.10.1093/icesjms/fst227

Singh NJ, Börger L, Dettki H, Bunnefeld N, Ericsson G (2012) From migration to nomadism: movement variability in a northern ungulate across its latitudinal range. Ecol Appl 22:2007–2020

Skud BE (1977) Drift, migration, and intermingling of Pacific halibut stocks. Int Pac Halibut Comm Sci Rep 63:1:42

Solmundsson J, Palsson J, Karlsson H (2005) Fidelity of mature Icelandic plaice (*Pleuronectes platessa*) to spawning and feeding grounds. ICES J Mar Sci 62:189-200

Starr RM, O'Connell V, Ralston S (2004) Movements of lingcod (*Ophiodon elongatus*) in southeast Alaska: potential for increased conservation and yield from marine reserves. Can J Fish Aquat Sci 61:1083-1094

Staudinger MD, Juanes F (2010) Feeding tactics of a behaviorally plastic predator, summer flounder (*Paralichthys dentatus*). J Sea Res 64:68-75

Stewart IJ, Leaman BM, Martell S, Webster RA (2013) Assessment of the Pacific halibut stock at the end of 2012. In: Sadorus L (ed) Report of Assessment and Research Activities 2012. IPHC, Seattle WA, p 93-186

Taggart SJ, Shirley TC, O'Clair CE, Mondragon J (2004) Dramatic increase in the relative abundance of large male dungeness crabs *Cancer magister* following closure of commercial fishing in Glacier Bay, Alaska. Am Fish Soc Symp 42:243 - 253

Tetreault I, Ambrose RF (2007) Temperate marine reserves enhance targeted but not untargeted fishes in multiple no-take MPAs. Ecol Appl 17:2251-2267

Tolimieri N, Andrews K, Williams G, Katz S, Levin PS (2009) Home range size and patterns of space use by lingcod, copper rockfish and quillback rockfish in relation to diel and tidal cycles. Mar Ecol Prog Ser 380:229-243

Trumble RJ, St-Pierre G, McGregor IR (1991) Evaluation of Pacific halibut management for regulatory area 2A. Int Pac Halibut Comm Sci Rep 74:1 - 44

Turchin P (1998) Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. Sinauer Associates, Inc., Sunderland, Massachusetts

Valero JL, Webster RA (2012) Current understanding of Pacific halibut migration patterns. In: Sadorus L (ed) Report of Assessment and Research Activities 2011. International Pacific Halibut Commission, Seattle, WA, p 341-379

Webster RA, Clark WG, Leaman BM, Forsberg JE (2013) Pacific halibut on the move: a renewed understanding of adult migration from a coastwide tagging study. Can J Fish Aquat Sci 70:642-653

Webster RA, Stewart IJ (2014) Apportionment and regulatory area harvest calculations. In: Sadorus L (ed) Report of Assessment and Research Activities 2013. IPHC, Seattle WA, p 197 - 216

Wood SN (2006) Generalized Additive Models: an introduction with R. Chapman and Hall/CRC, Boca Raton, FL

Wright DJ, Lundblad ER, Larkin EM, Rinehart RW, Murphy J, Cary-Kothera L, Draganov K (2005) ArcGIS Benthic Terrain Modeler: a collection of tools used with bathymetric data sets to examine the deepwater benthic environment. Oregon State University, Davey Jones' Locker Seafloor Mapping/Marine GIS Laboratory and NOAA

Table 1.1. Tagged fish ID, release date, total length, maximum horizontal displacement observed, net horizontal displacement at last observation, dispersal pattern model code (see Table 1.2), tracking year, date of each observation, and movement state code (see legend).

Fish ID	Release Date	Length (cm)	Max disp. (m)	Net disp. (m)	Model code	Tracking year
1	7/24/91	136	13334	6369	FSHR	1991
2	7/30/91	91	2172	1475	HR	
3	8/20/91	136	1492	150	HR	
4	8/20/91	115	-	-	N	
5	8/20/91	140	9796	9614	SHR	
6	8/20/91	78	1778	984	HR	
7	8/20/91	90	2031	319	HR	
8	9/9/91	96	14177	14177	R	
9	9/9/91	117	9619	8023	SHR	
10	6/3/92	157	3691	2027	FSHR	1992
11	6/17/92	128	16658	11472	U	
12	6/19/92	94	-	-	N	
13	6/19/92	68	4558	4558	U	
14	6/19/92	76.5	762	762	U	
15	6/23/92	153	874	378	HR	
16	7/1/92	129	9291	418	SF	
17	7/1/92	104	2390	2390	R	
18	7/8/92	146	2269	1497	HR	
19	7/8/92	127	2163	2163	U	
20	7/8/92	132.5	13393	13393	SHR	
21	7/8/92	121	1848	213	HR	
22	7/8/92	124	8184	1909	SF	
23	7/8/92	142.5	237	237	U	
24	7/8/92	134	16478	16478	R	
25	7/8/92	125	12616	3993	SF	
26	7/8/92	129	12494	3839	SF	
27	7/31/92	201.3	2014	1296	HR	
28	11/17/92	130.6	9188	207	SF	1993
29	11/17/92	149	10872	496	SF	
30	11/17/92	132.5	1652	497	U	
31	11/17/92	168	614	73	HR	
32	6/2/93	220	3190	1949	HR	
33	6/18/93	154	17954	15963	R	
34	6/18/93	162	-	-	N	
35	6/18/93	151.2	-	-	N	
36	6/19/93	140	1170	567	HR	
37	6/19/93	182.8	3526	2614	HR	
38	6/19/93	155	859	843	HR	
39	6/19/93	151.6	-	-	N	
40	6/20/93	152.8	2269	507	HR	
41	6/20/93	149.8	1109	726	HR	
42	6/20/93	164.4	1296	359	HR	
43	8/11/93	77.6	807	619	HR	

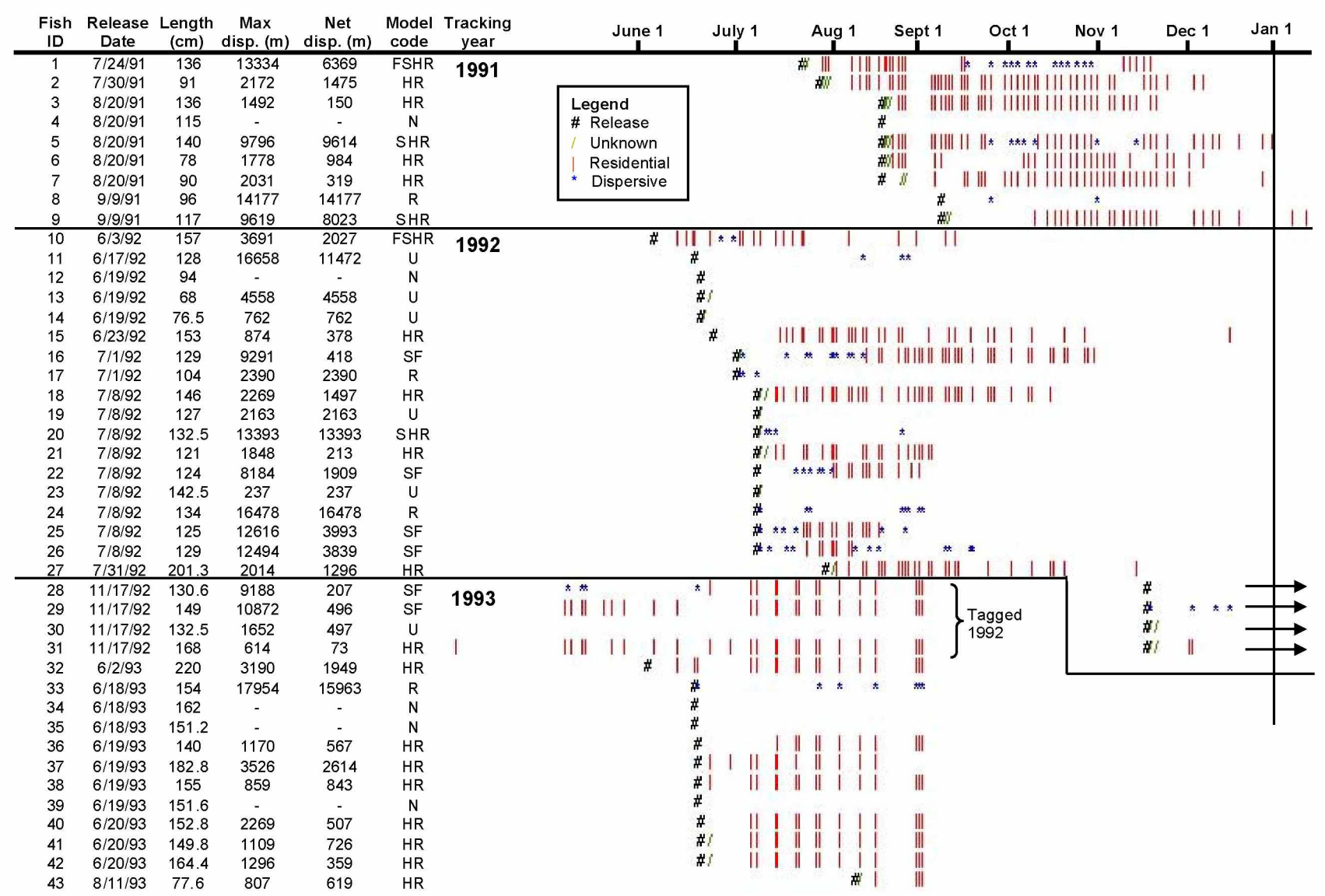


Table 1.2. Model name, code, number of parameters, theoretical example, and formula for dispersal patterns used to classify movement trajectories.

Model (code)	No. params	Ecological interpretation	NSD vs. time	Formula
1. Home range (HR)	1*	Constrained movement, such as home range behavior (residential movement state)		$NSD = \propto$
2. Random (R)	1	Nomadic, diffusive movement such as foraging		$NSD = \beta t$
3. Shifted home range (SHR)	3	Movement to new location		$NSD = \frac{\delta}{1+\exp\left(\frac{\theta - t}{\phi}\right)}$
4. Site fidelity (SF)	4	Migration or foray with return to original location		$NSD = \frac{\delta}{1+\exp\left(\frac{\theta_m - t}{\phi}\right)} + \frac{-\delta}{1+\exp\left(\frac{\theta_r - t}{\phi}\right)}$
5. Foray/shifted home range (FSHR)	6	Migration or foray with return to new location	 Time	$NSD = \frac{\delta_m}{1+\exp\left(\frac{\theta_m - t}{\phi_m}\right)} + \frac{-\delta_r}{1+\exp\left(\frac{\theta_r - t}{\phi_r}\right)}$

Where t is time, $\propto$ is the intercept of a linear model, β is the slope of a linear model, and δ, θ, and φ are non-linear model coefficients that describe the spatial extent, timing, and temporal scale respectively. Subscript "m" refers to "migration start", "r" refers to return. See Bunnefeld et al. (2011). *Home range can be modeled as intercept-only when the sampling interval is greater than the time the animal takes to reach the limits of its range.

Table 1.3. *Hippoglossus stenolepis.* Number of animals assigned to each dispersal pattern (see Table 1.2; HR = home range, SHR = shifted home range, FSHR= foray/shifted home range, SF = site fidelity, R = random, U = unknown, N = not relocated after release) along with average maximum horizontal displacement, average net displacement at the end of the observation period, and average observation period duration for each pattern (SE).

Model	No. Animals	Avg. Max displacement (km)	Avg. Net displacement (km)	Avg. Obs. Duration (days)	Avg. total length (cm)
HR	17	1.7 (0.02)	0.9 (0.2)	102.5 (14.3)	142.7 (10.0)
SHR	3	10.9 (1.2)	10.3 (1.6)	101.7 (26.4)	129.8 (6.8)
FSHR	2	8.5 (4.8)	4.2 (2.2)	110.0 (8.0)	146.5 (10.5)
SF	6	10.4 (0.8)	1.8 (0.7)	146.5 (46.5)	131.1 (3.7)
R	4	12.8 (3.5)	12.3 (3.3)	48.3 (14.8)	122.0 (13.4)
U	6	4.3 (2.5)	3.3 (1.8)	13.5 (11.7)	134.8 (12.9)
N	5	NA	NA	NA	112.4 (12.9)

Table 1.4. *Hippoglossus stenolepis.* Habitat associations and habitat occupancy ranges for the residential movement state. Significant Kolmogorov-Smirnov test statistic D in bold text, one-tail test sign (positive = the upper 95% C.I. for the observed CDF is greater than the available CDF, negative = the lower 95% C.I. for the observed CDF is less than the available CDF), p values, value of habitat where significant D+ or D- statistics were obtained.

Habitat variable	D	p value	One-tail test sign	Habitat value at D
Depth (m)	**0.1701**	0.003	+	104
Slope (deg)	**0.2674**	< 0.001	-	1.0
Change in slope (deg)	**0.3234**	< 0.001	-	3.6
Rugosity (ratio)	**0.2066**	< 0.001	-	0.00015
Tide speed (m/s)	**0.2114**	< 0.001	-	0.07
Tide speed (m/s)	**0.2189**	< 0.001	+	0.34
Distance to soft bottom (m)	0.0127	0.968	+	N.S.
Distance to moderate complexity (m)	0.0785	0.288	+	N.S.

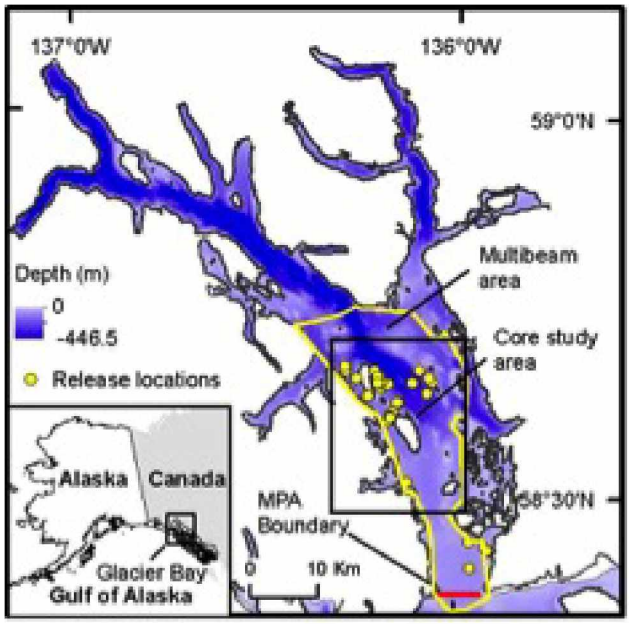


Figure 1.1. Location of the study area within the inside waters of Glacier Bay National Park. Map depicts the MPA boundary (red line), multibeam survey area for habitat analyses (yellow line), tagged animal release locations (yellow circles), and core study area (black square) where most tagged animals were released and tracked.

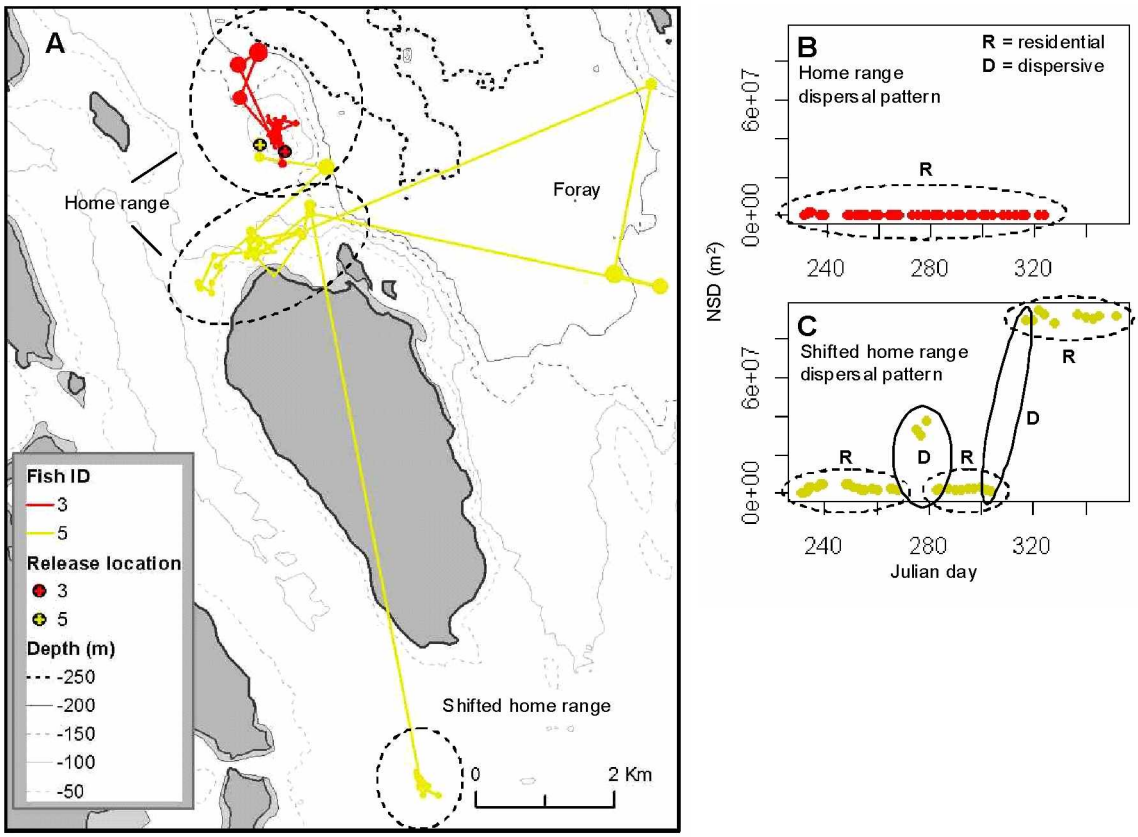


Figure 1.2. *Hippoglossus stenolepis.* A) Locations for two tagged fish (#3 and #5) with different dispersal patterns. The size of each circle represents estimated telemetry position error (in meters). Corresponding plots of NSD vs. time for a B) home range dispersal pattern with all observations classified as residential movement (tag #3), and C) a shifted home range dispersal pattern with observations classified as residential (dashed line ovals) or dispersive (solid line ovals) movement states (tag #5).

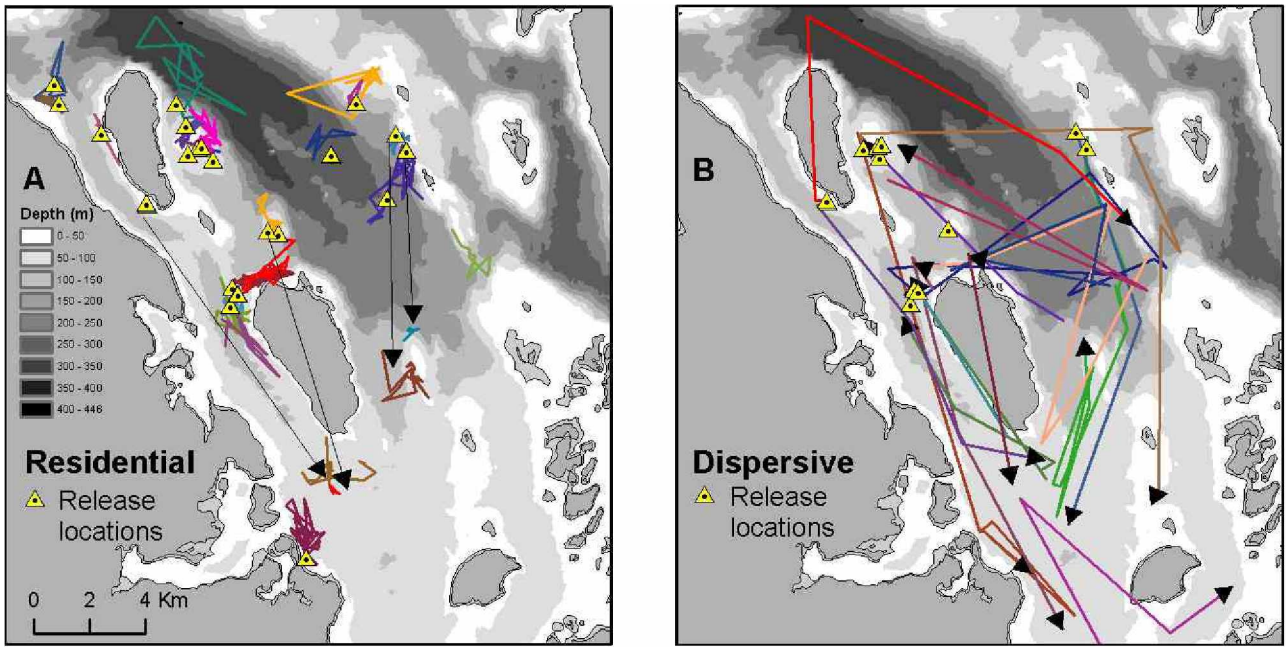


Figure 1.3. *Hippoglossus stenolepis.* Residential and dispersive movement within the core study area. A) Release locations for all animals that exhibited residential movement are shown by yellow triangles, residential movement sequences are shown by colored lines, and shifted home ranges from release locations are indicated by black arrows. B) Release locations for all animals that exhibited dispersive movement are shown by yellow triangles and dispersive movement sequences are shown by colored arrows.

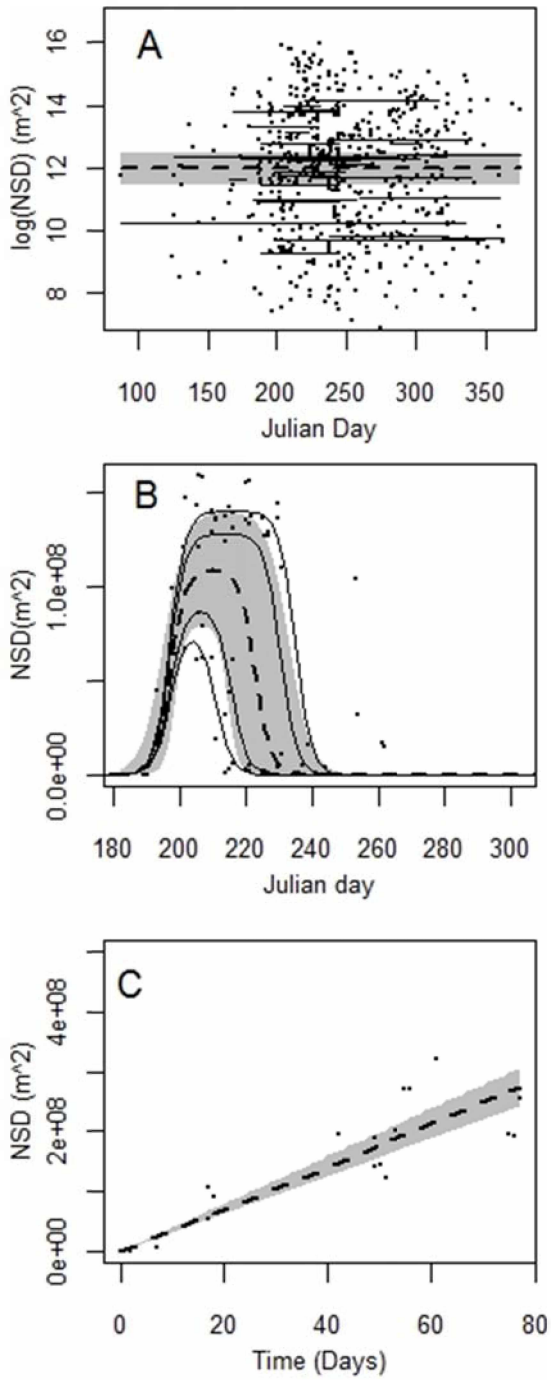


Figure 1.4. *Hippoglossus stenolepis.* Observed (points) and estimated (lines) NSD from mixed-effects models for A) the residential movement state (n = 31 home ranges) on log-scale, B) the site fidelity dispersal pattern (n = 4 animals, as 2 animals with winter observations were not included), and C) the random dispersal pattern (n = 4 animals). Population (fixed-effect) means are shown with thick dashed lines, 95% C.I. are shown as gray polygons, and thin black lines represent individual (random-effects) estimates. In (A), the length of the thin lines represents the period of observation for each animal.

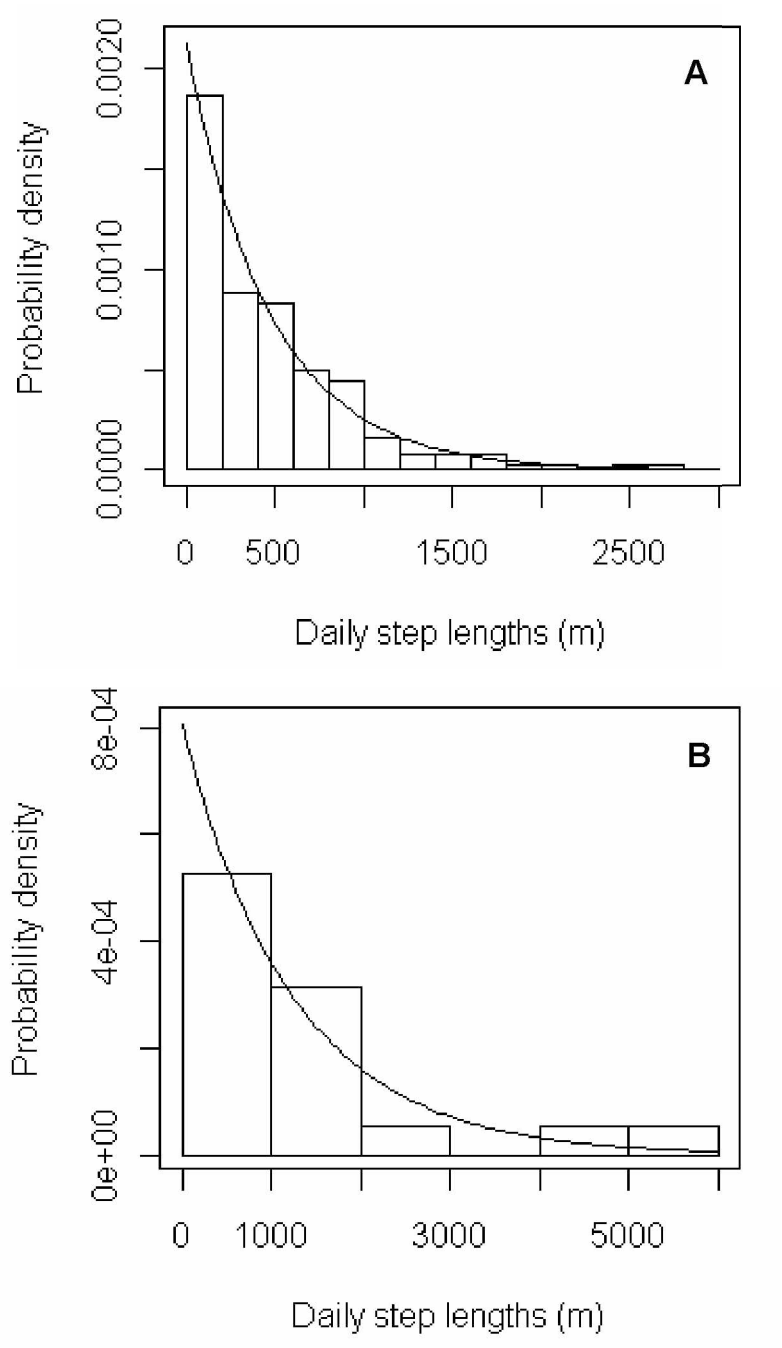


Figure 1.5. *Hippoglossus stenolepis.* Daily step length distributions for A) residential and B) dispersive movement states along with exponential curves fitted to each distribution.

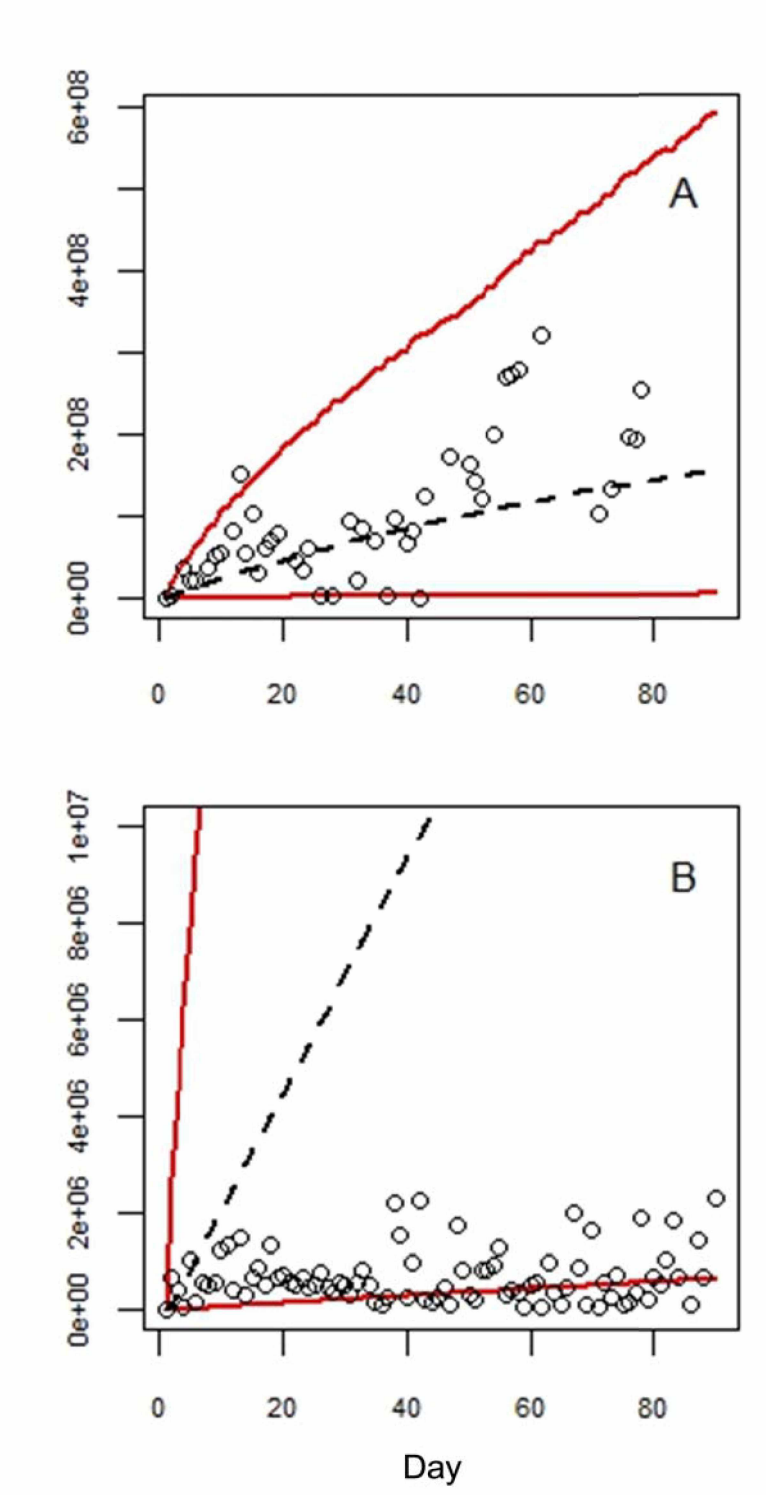


Figure 1.6. *Hippoglossus stenolepis.* Observed mean NSD (circles) compared to correlated random walk simulations (lines) for A) dispersive and B) residential movement states. Note the difference in scales for the y axes. The average mean value from 1000 simulations is shown with a dashed black line, and average upper and lower 95% C.I.s are shown with red lines.

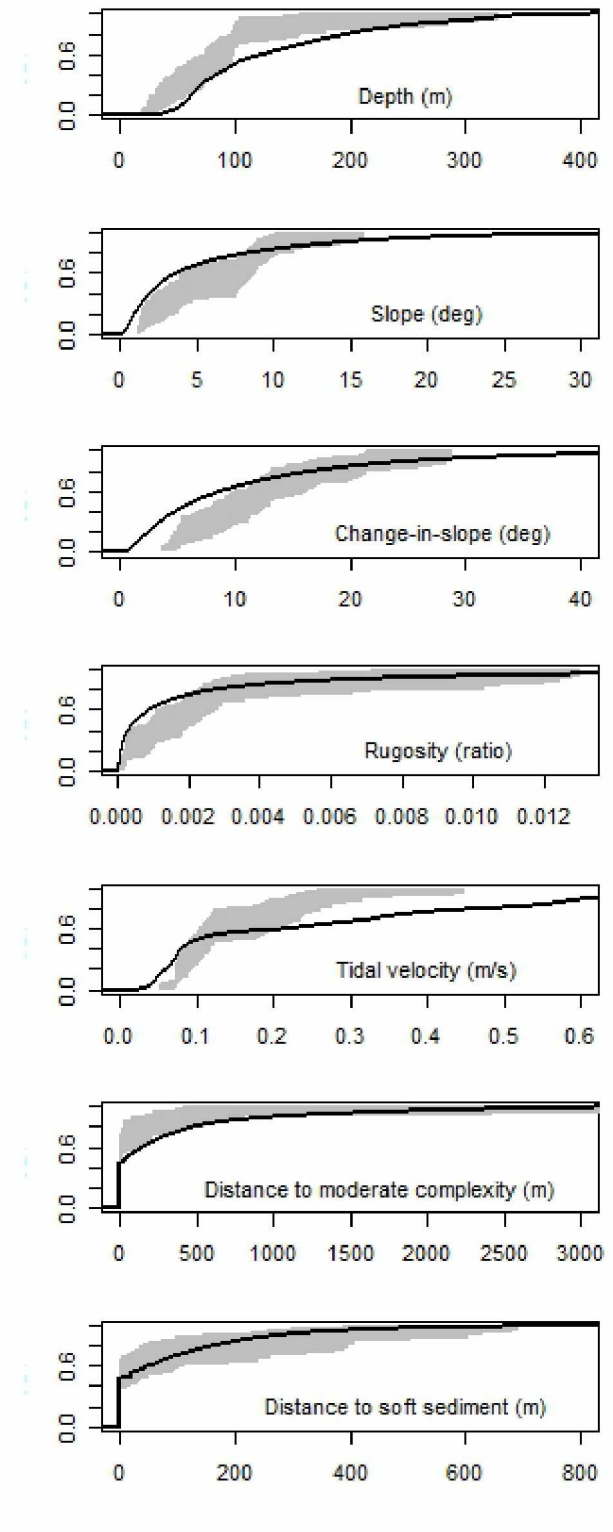


Figure 1.7. *Hippoglossus stenolepis.* Observed (polygons) and available (thick black line) cumulative distribution functions (CDFs) for depth, slope, change-in-slope, rugosity, tidal velocity, distance to moderately complex habitat, and distance to soft substrate habitat. 95% C.I.s for residential movement CDFs are indicated by gray polygons. See Table 1.4 for habitat association test statistics.

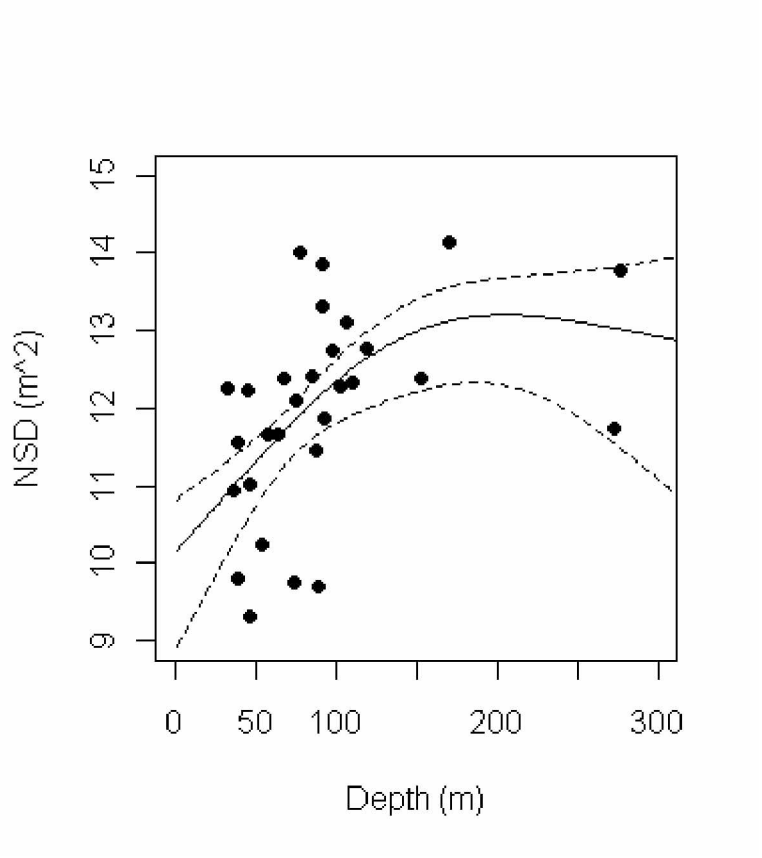


Figure 1.8. *Hippoglossus stenolepis.* Predicted relationship between depth and home range size from best GAM (solid black line); 95% C.I. are shown in dashed lines.

1.8 Appendices

1.8.1 Supplement 1-1: Fish tagging and tracking

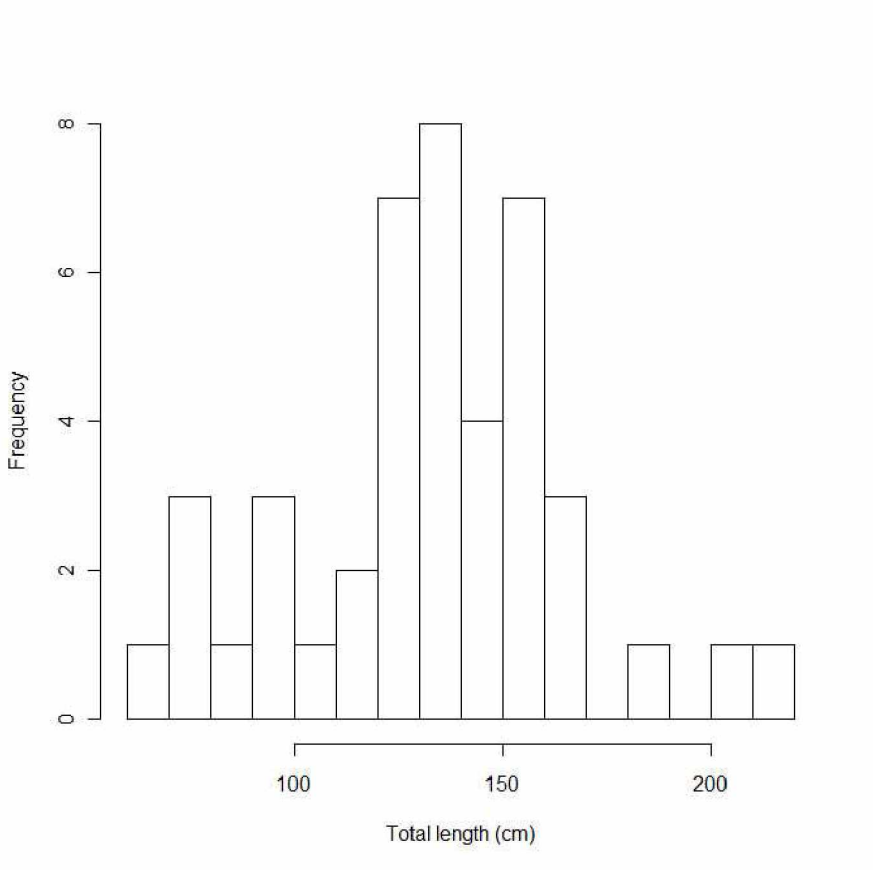


Figure S1.1-1. *Hippoglossus stenolepis*. Length-frequency distribution of fish tagged and released with acoustic tags.

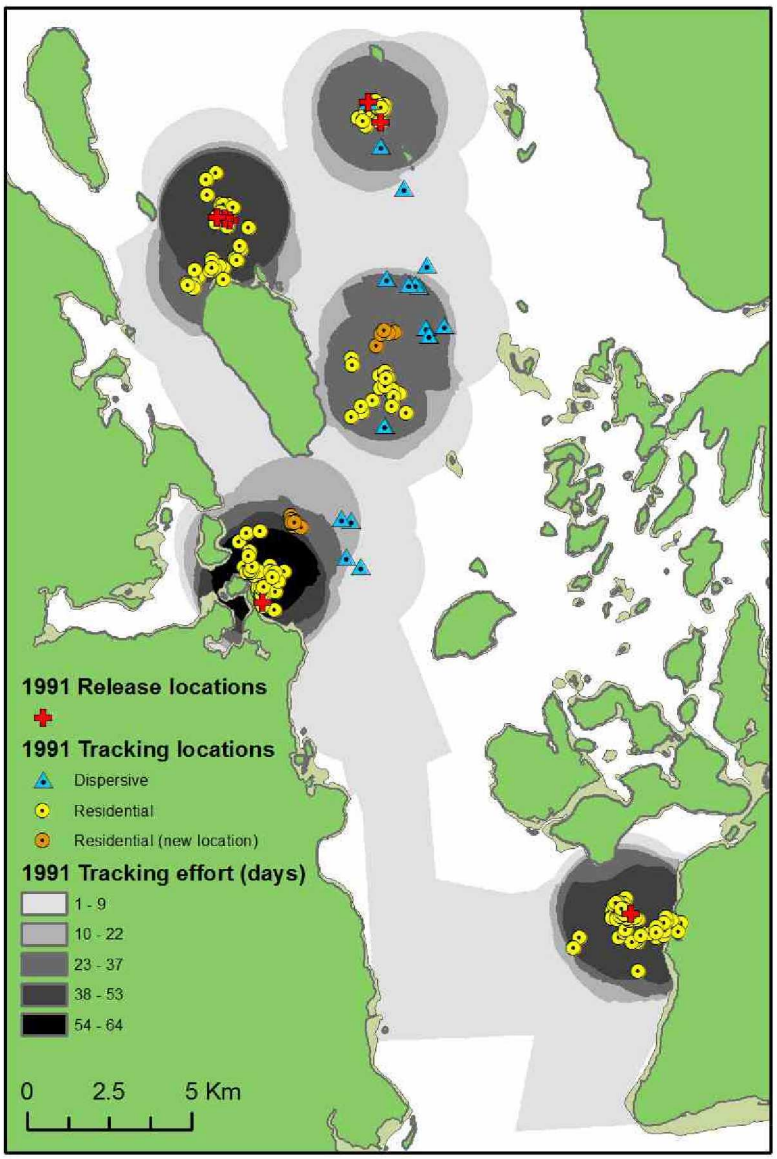
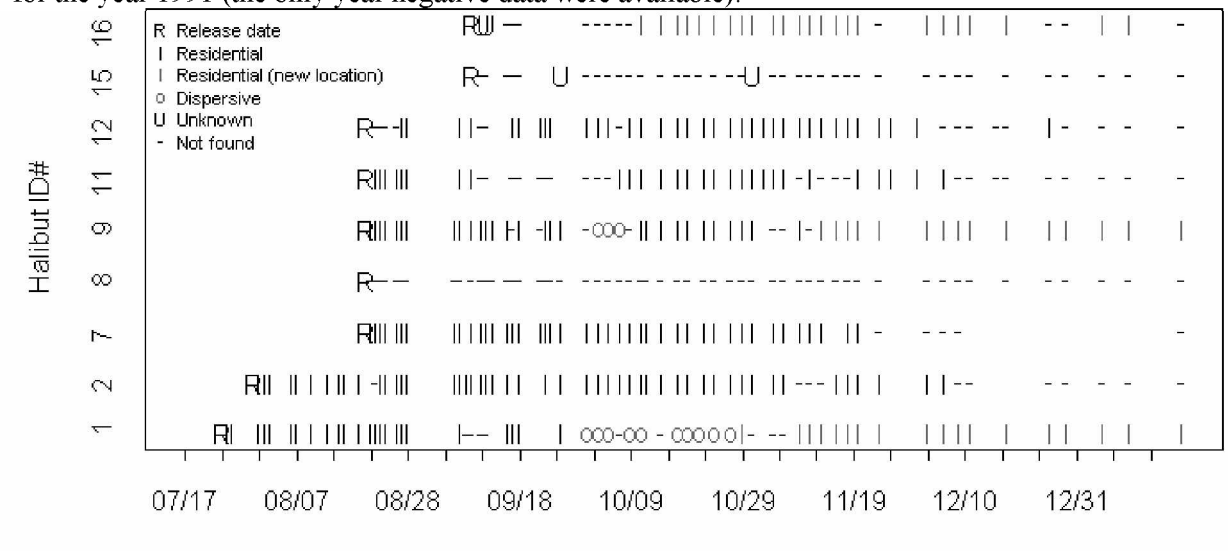


Figure S1.1-2. *Hippoglossus stenolepis.* Grid of tracking effort (days searched) during 1991 along with 1991 release locations and positions of tagged animals in the dispersive movement state (blue triangles), the residential movement state (yellow circles), and the residential movement state in new locations (e.g., shifted home ranges; orange circles).

Table S1.1-1. Acoustic tag specifications (signal frequency and size) and attachment method (external or surgical implantation) for individual tagged animals. Small tags were 95 mm long x 18 mm diameter and weighed 16 g in water. Large tags were 95 mm long x 34 mm diameter and weighed 34 g in water.

ID	Release date	Frequency (Khz)	Tag size	Attachment method
1	24-Jul-91	80	Large	External
2	30-Jul-91	80	Small	External
3	20-Aug-91	80	Large	External
4	20-Aug-91	80	Large	External
5	20-Aug-91	80	Large	External
6	20-Aug-91	80	Small	External
7	20-Aug-91	80	Small	External
8	09-Sep-91	80	Small	External
9	09-Sep-91	80	Large	External
10	03-Jun-92	35	Large	External
11	17-Jun-92	35	Small	External
12	19-Jun-92	35	Large	External
13	19-Jun-92	35	Small	External
14	19-Jun-92	35	Small	External
15	23-Jun-92	35	Large	External
16	01-Jul-92	35	Large	External
17	01-Jul-92	35	Small	External
18	08-Jul-92	35	Large	External
19	08-Jul-92	35	Small	External
20	08-Jul-92	35	Large	External
21	08-Jul-92	35	Large	External
22	08-Jul-92	35	Small	External
23	08-Jul-92	35	Large	External
24	08-Jul-92	35	Small	External
25	08-Jul-92	35	Large	External
26	08-Jul-92	35	Small	External
27	31-Jul-92	35	Small	Internal
28	17-Nov-92	35	Large	Internal
29	17-Nov-92	35	Large	Internal
30	17-Nov-92	35	Small	Internal
31	17-Nov-92	35	Small	Internal
32	02-Jun-93	35	Small	Internal
33	18-Jun-93	35	Large	Internal
34	18-Jun-93	35	Large	Internal
35	18-Jun-93	35	Large	Internal
36	19-Jun-93	35	Large	Internal
37	19-Jun-93	35	Large	Internal
38	19-Jun-93	35	Large	Internal
39	19-Jun-93	35	Large	Internal
40	20-Jun-93	35	Large	Internal
41	20-Jun-93	35	Large	Internal
42	20-Jun-93	35	Large	Internal
43	11-Aug-93	35	Small	Internal

Table S1.1-2. *Hippoglossus stenolepis.* Information on movement states (residential, dispersive, or unknown) along with information on whether an animal was searched for and not found on a given day for the year 1991 (the only year negative data were available).



1.8.2 Supplement 1-2: Model selection examples

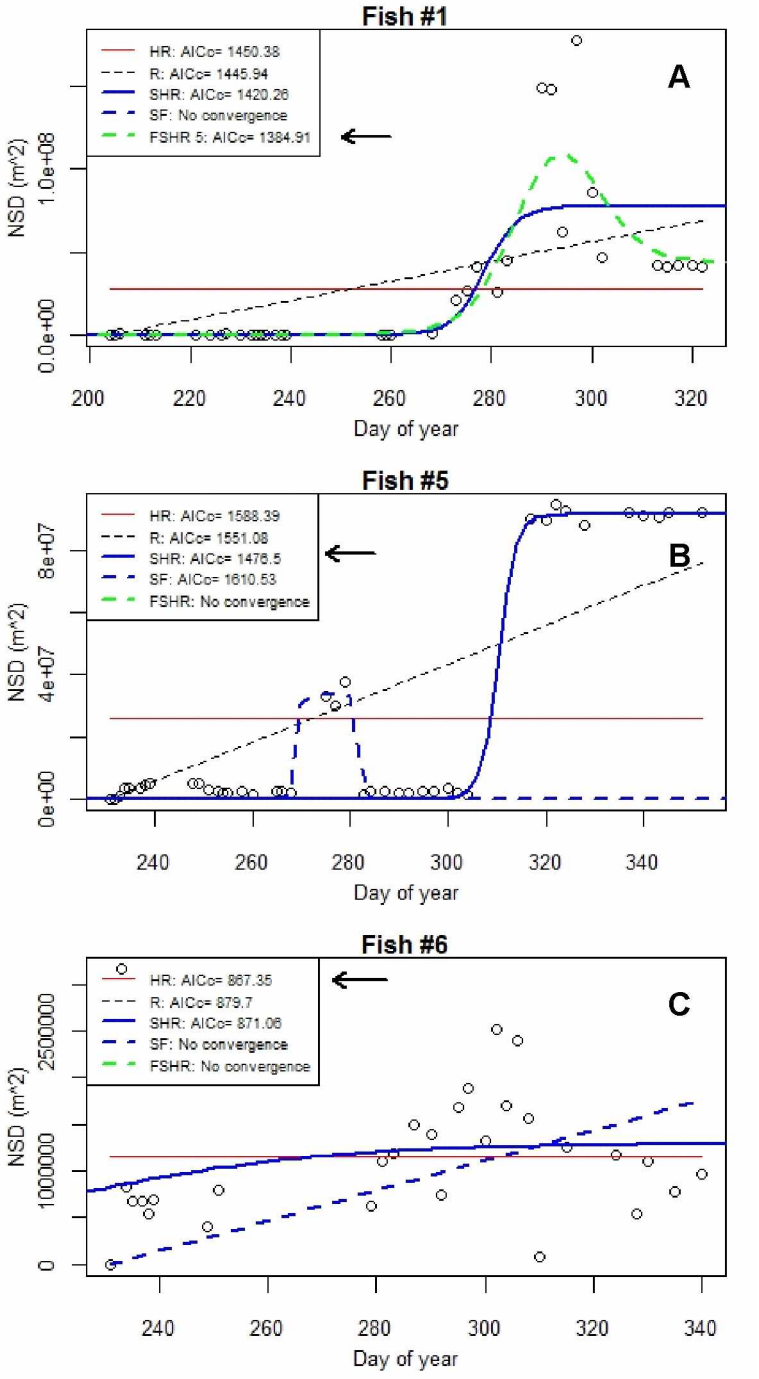


Figure S1.2-1. *Hippoglossus stenolepis.* Classification of (A) foray/shifted home range (FSHR), (B) shifted home range (SHR), and (C) home range (HR) dispersal patterns for individual fish based on model selection with AICc (see Table 1.2 for model descriptions). AICc values are the lowest for the best-fitting model, shown by the arrows for each fish.

1.8.3 Supplement 1-3: Habitat

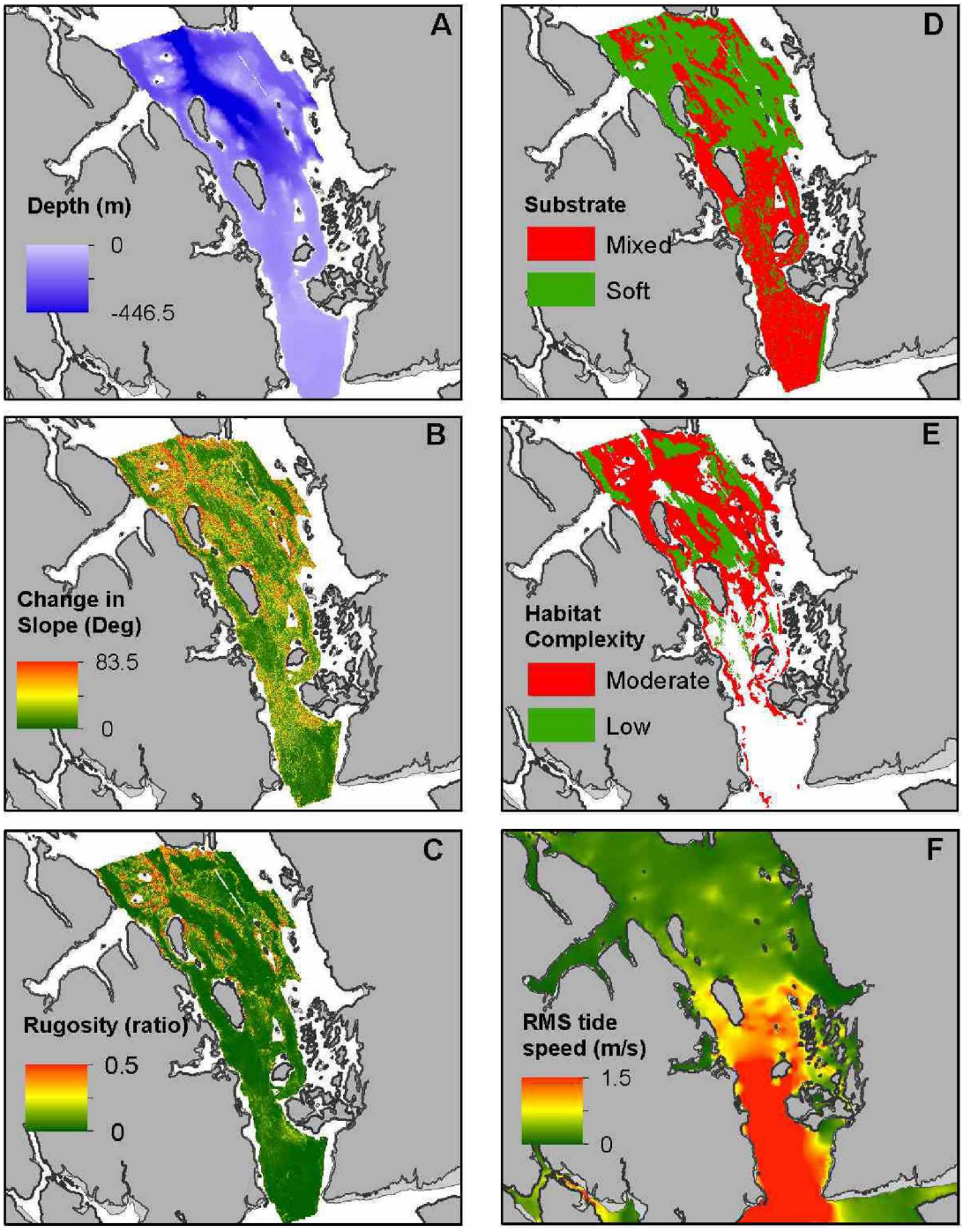


Figure S1.3-1. Environmental information used for habitat analyses: A) depth, B) change in slope (slope not pictured), C) rugosity, D) substrate type polygons, E) habitat complexity polygons, and F) RMS tidal velocity.

Table S1.3-1. Detailed information on continuous habitat raster resolution, range of values in 20 m grid of study area (n = 1.08E06), and habitat variable source/derivation.

Habitat Variable	Raster resolution (m)	Range	Source
Depth (m)	5	10 - 413	Multibeam bathymetry (Carlson et al. 2002)
Slope (deg)	5	0 - 83.7	Derived from bathymetry raster using ArcGIS 10.0 Spatial Analyst slope tool
Change in slope (deg)	5	0 - 83.1	Derived from slope raster using ArcGIS 10.0 Spatial Analyst slope tool
Rugosity (ratio)	5	2.0E-06 - 0.498	Derived from bathymetry raster using the Vector Ruggedness Measure command in Benthic Terrain Modeler ArcGIS10.1 (Wright et al. 2005)
RMS tidal velocity (m/s)	100	1.11E-02 - 0.254	Output from circulation model (ADCIRC) for Glacier Bay (David Hill, OSU)
Distance from soft substrate polygons (m)	20	0 - 1825	Derived from Glacier Bay Habitat Map (Harney et al. 2006) using ArcGIS10.0 Euclidean distance tool
Distance from moderate habitat complexity polygons (m)	20	0 - 3617	Derived from Glacier Bay Habitat Map (Harney et al. 2006) using ArcGIS10.0 Euclidean distance tool

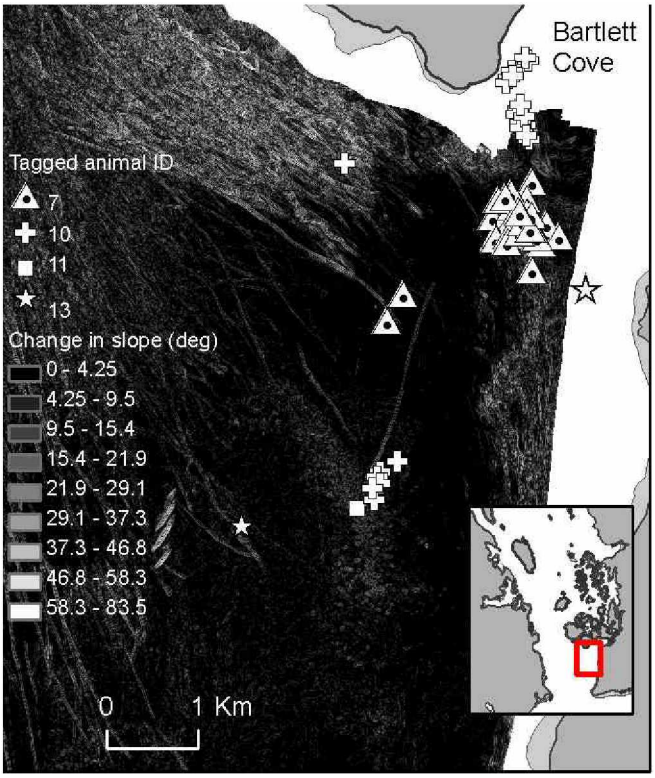


Figure S1.3-2. *Hippoglossus stenolepis.* Example of tagged animal association with the change in slope habitat variable. Tagged animal observations for 4 animals in close proximity to areas with high heterogeneity (shown in light gray colors). Glacial scours and moraine are apparent in otherwise flat terrain.

Chapter 2. Interannual site fidelity of Pacific halibut: potential utility of protected areas for management of a migratory demersal fish[1]

2.1 Abstract

Pacific halibut is a large-bodied demersal fish species known to undertake large-scale winter spawning migrations. We characterized annual movement patterns of Pacific halibut relative to a Marine Protected Area (MPA) in Glacier Bay National Park, Alaska using electronic tags. In the summer of 2013 we deployed 25 Pop-up Satellite Archival Tags (PSATs) on halibut to determine seasonal movement patterns, where residency within the MPA was inferred by comparing PSAT depth and temperature records to stationary archival tag data and oceanographic survey data. To characterize within-summer movement patterns, 15 halibut tagged with PSATs were double-tagged with acoustic transmitters and actively tracked during the summers of 2013 and 2014. Home range behavior and interannual site fidelity at spatial scales < 5 km were observed for the majority of tagged fish. A small proportion (6/21 fish with PSAT data) departed the MPA on winter migrations during December 2013. A majority (4/6) of migratory fish returned to the MPA after an average of 57 days (s.d. 22 days) spent outside of the MPA. Migration timing generally coincided with existing winter commercial fishery closures. The annual movement patterns of tagged halibut relative to MPA boundaries and winter commercial fisheries closures suggest that the Glacier Bay MPA could serve as a year-round refuge from commercial harvest for both residential and migratory halibut. If halibut behavior is similar in other areas, protected areas may provide some utility for management despite the migratory nature of halibut.

2.2 Introduction

Pacific halibut, *Hippoglossus stenolepis* (subsequently referred to as "halibut"), is the subject of important commercial, sport, and subsistence fisheries in the North Pacific Ocean. The estimated female spawning biomass has declined by more than 50% over the last 20 years (Stewart et al., 2016). This

[1] This is a pre-copyedited, author-produced version of an article accepted for publication in ICES Journal of Marine Science following peer review. The version of record [Nielsen, J. K., and Seitz, A. C. 2017. Interannual site fidelity of Pacific halibut: potential utility of protected areas for management of a migratory demersal fish. ICES J Mar Sci, 74: 2120–2134] is available online at: https://doi.org/10.1093/icesjms/fsx040.

decrease stems from reductions in size at age, particularly for older adult females, along with declines in numbers of fish (Stewart and Monnahan, 2016). This decline has resulted in one of the region's most pressing fisheries management concerns, as the allocation of the total available yield to different users becomes more challenging when the halibut biomass is low.

Halibut is managed as a panmictic population due to large-scale dispersal at the larval and juvenile life stages (Skud, 1977; Clark and Hare, 2006; Valero and Webster, 2012). Adults are also known to move large distances (Webster et al., 2013), particularly during spawning migrations that occur when they move from inshore waters occupied during summer to offshore waters to spawn during winter (Loher and Seitz, 2006). Recent electronic tagging studies have demonstrated a tendency to return to previously-occupied summer feeding grounds after undertaking seasonal spawning migrations (Loher, 2008). These observations suggest that halibut may exhibit site fidelity to summer foraging locations, and thus some degree of spatial structure may be present among adult halibut.

Protected areas may have some utility for management of migratory marine species that have spatial structure at some stage in their life cycles, such as site fidelity to feeding or spawning locations (Lascelles et al., 2014). Knowledge of how fish move in relation to MPA boundaries is critical for implementing MPAs that are designed for achieving specific goals, such as increasing abundance and size structure within reserves or "spillover" of juveniles and adults to areas adjacent to MPAs (Grüss et al., 2011). To evaluate whether such benefits could also occur for mobile marine species such as halibut, information on movement characteristics including the timing and proportion of individual animals that leave the MPA is needed.

Glacier Bay National Park, a fjord in Southeast Alaska, U.S.A., provides an opportunity to study the potential benefits of MPAs for halibut. Although both commercial and sport fishing are currently allowed in Glacier Bay National Park, regulations exist that limit both types of harvest in the interior fjord portion of its marine waters known as "Glacier Bay Proper." Legislation enacted in 1998 (36 CFR 13.1130-1146) initiated a phase-out of the commercial fishery for halibut in Glacier Bay Proper, subsequently referred to as "the MPA", that is estimated to be completed sometime after 2050. This MPA

is one of only three areas where the commercial harvest of halibut is restricted in the state of Alaska (Witherell and Woodby, 2005). Sport fishing within the MPA will continue after the commercial fishery is phased out. However, sport fishing is restricted during the summer months (June through August) by daily vessel entry limits of 6 charter and 25 private vessels (36 CFR 13.1160).

Previous research on fine-scale movement of adult halibut using acoustic telemetry in the MPA indicated that most tagged halibut remained in the MPA during the summer and occupied home ranges at scales of < 2 km (Nielsen et al., 2014, Chapter 1). This was an important finding for understanding the potential effectiveness of the MPA after commercial fishing ceases, as certain benefits such as increased abundance or size structure may be achieved within protected areas that encompass the range of fish movement (Halpern, 2003; Edgar et al., 2014). However, information on the movement of Glacier Bay halibut during winter, when adults are thought to undertake migrations to spawning areas, was not available.

The goal of this study is to assess the potential effectiveness of the MPA for protecting halibut spawning stock by providing information on annual movement patterns in relation to MPA boundaries and seasonal fisheries closures. We present results of adult halibut movement based on a combination of Pop-up Satellite Archival Tags (PSATs) and acoustic telemetry to determine movement at scales of hundreds of m within the MPA as well as detect movement outside of the MPA (> 20 km). The results are interpreted in terms of spatial and temporal scales of site fidelity, migration characteristics in fjord systems, and implications of annual movement patterns for the utility of protected areas for halibut management.

2.3 Methods

2.3.1 Study area

The study was conducted in the interior fjord portion of Glacier Bay National Park, located in northern Southeast Alaska (Figure 2.1A). Glacier Bay is a large, recently de-glaciated fjord-type estuarine system with depths to 450 m in the fjord arms. A glacial sill with a depth of approximately 40 m lies just outside the regulatory boundary at the bay's entrance (Figure 2.1B).

Glacier Bay National Park was designated a national monument in 1925 and achieved National Park status through congressional designation with the Alaska National Interest Lands Conservation Act (ANILCA) in 1980. In 2009, Glacier Bay National Park was officially designated an MPA in alignment with Executive Order 13158 and joined the U.S. national system of MPAs. The technical boundary of the MPA established in 2009 includes the interior fjord portion, known as Glacier Bay Proper, as well as outside waters that extend into Icy Strait and three miles into the Gulf of Alaska from the southern and western shores of Glacier Bay National Park. Because the harvest of halibut is only limited by National Park Service regulations in the interior fjord (36 CFR 13.1130-1146), we refer to this area as the functional MPA for halibut and frame our analysis on MPA residency and migration of halibut relative to its boundary at the entrance to the fjord (Figure 2.1B). Throughout this document, the term "MPA" refers to the functional MPA in the interior fjord rather than the official MPA established in 2009. Fish were tagged and tracked in a 25 km^2 core study area in the middle of the MPA (Figure 2.1C).

2.3.2 Electronic tags

To understand movement of adult halibut in Glacier Bay at multiple spatial and temporal scales, two types of electronic tags were attached to fish during the summer of 2013. Pop-up Satellite Archival Tags (PSATs) were deployed to provide data on migratory characteristics at monthly to yearly timescales. Acoustic transmitters were deployed on a subset of PSAT-tagged fish to characterize horizontal movements during summer and to obtain fine-scale information on dispersal distances within one year after tagging.

2.3.3 Capture and tagging

Halibut were captured on 13 longline sets in the core study area (Figure 2.1C) between June 29 and July 2, 2013. Each set was configured with approximately 100 snap-on 16-0 circle hooks baited with pink salmon (*Oncorhynchus gorbuscha*). Soak times averaged 12.2 hours (range 1.6 to 23.7 hours). Halibut that met minimum size requirements (≥105 cm total length) were carefully lifted on board and subdued by placing a wet towel over their eyes while post-release viability was assessed and measurements of total length were obtained. Only fish that displayed normal coloration and opercular

movements and did not display frayed fins or bleeding wounds were tagged. PSATs with 12.7-cm tethers of 136-kg-test nylon monofilament line were attached by inserting a surgical-grade titanium dart through the skin and pterygiophores (fin ray supports), roughly 2.5 cm medially from the dorsal fin on the eyed-side of the halibut where the body began to taper towards the tail (Seitz et al., 2003).

Acoustic transmitters were surgically implanted in the peritoneal cavity. Tags were placed in autoclave envelopes and sterilized with ethylene oxide prior to deployment. Separate sterilized packs of surgical tools were used for each fish. A 2-cm incision was made through the peritoneal wall roughly parallel to the long axis of the fish, positioned about 1 cm dorsal to the anterior end of the first interhaemal spine. The tag was then inserted into the peritoneum and the entry incision was closed using three to five sutures of 2-0 prolene suture material with a reverse cutting FS needle. Once tagged, fish were gently released head-first back into the ocean. Time on board, depth at release, and position at release were recorded for each tagged animal. Research was conducted under permits from the University of Alaska Fairbanks Institutional Animal Care and Use Committee (permit number 215255), International Pacific Halibut Commission (IPHC), and the National Park Service (permit number GLBA-2013-SCI-0017).

2.3.4 Pop-up satellite archival tags (PSATs)

PSATs continuously record environmental data such as depth and temperature while externally attached to a fish. On a specified, pre-programmed date, the PSAT releases from the animal, floats to the ocean surface, and transmits its recorded data to the Argos satellite network. Estimated positions of the floating tag are determined from successive uplinks to Argos satellites along with error estimates that can range from less than 250 m to greater than 1500 m. The most precise position of the tag obtained within the first 1 – 2 hours after the tag reaches the surface is assumed to represent the location of the tagged fish on the last day that archival data were recorded and is therefore referred to as the "end location". Archived data retrieved from the tag can be used to infer the location of the tagged fish on each day that archival data were collected by matching the tag data to modeled or measured values in the study area, a process known as "geolocation" (Evans and Arnold, 2009).

We deployed PSATs from two different manufacturers on a total of 25 halibut. Fifteen fish were tagged with Desert Star Systems (Marina, CA, U.S.A., subsequently referred to as "DS") SeaTag MOD archival tags that weighed 163 g in air with a length of 275 mm, a diameter of 25 mm for the tube, and a maximum diameter of 54 mm for the float. These tags recorded a detailed time series of depth (resolution 1 m, offset accuracy ± 20 m), temperature (resolution 0.002°C, accuracy ± 0.1°C), light intensity, and other variables at a 4-minute measurement interval. Archived data were summarized for transmission to the Argos satellite network by calculating the maximum depth, minimum depth, average depth, and average temperature. Ten fish were tagged with Wildlife Computers (Redmond, WA, U.S.A., subsequently referred to as "WC") MK-10 tags that weighed 75 g in air with a length of 175 mm, a diameter of 21 mm for the tube, and a maximum width of 40 mm for the float. These tags recorded depth (resolution 0.5 m, offset accuracy ± 10 m), temperature (resolution 0.05°C, accuracy ± 0.1°C), and light intensity at 10 second intervals. Data were summarized for transmission to satellites in 12-hour bins (beginning at 6:00 a.m. and 6:00 p.m. AKDT each day). Summarized data consisted of the maximum and minimum depth and temperature-depth profiles (maximum and minimum temperatures from eight depth bins that evenly spanned the range in depths observed for each time bin). For analyses, data from 12-hour bins were combined into daily 24-hour bins. Complete, detailed data sets (measurement intervals of 10 s for WC or 4 min for DS) were available if PSATs were physically recovered. Only depth and temperature records from both types of tags were used for analyses.

PSATs were programmed to release from the fish on either a winter or a summer pop-up date. Fifteen PSATs (8 DS, 7 WC) were programmed to release from the fish and transmit summarized data to satellites on February 1, 2014 to determine whether the end locations for tagged fish were inside or outside the MPA during the winter spawning season (Loher, 2011). Ten PSATs (7 DS, 3 WC) were programmed to release from the fish on July 1, 2014 to determine whether the end locations for tagged fish were located inside or outside of the MPA one full year after tagging. In some cases, end locations were not available even though PSAT data were obtained. For example, several DS tags did not begin to

transmit data until several weeks after the programmed pop-up date, but because data collection ceased on the pop-up date, the location of the transmitting tag could not be linked to the location of the fish at the end of the archival record.

2.3.5 Acoustic transmitters

Acoustic transmitters were also deployed to obtain position estimates of tagged fish present within the detection range of hydrophones towed by a vessel in the study area during the summer months. These position estimates were used to characterize horizontal movements of tagged fish during summer and provide independent locations of PSAT-tagged fish. Each of the fifteen fish that were tagged with a Desert Star PSAT were also tagged with a Lotek Wireless (Newmarket, Ontario, Canada) MAP acoustic telemetry transmitter (MM-M-16-50). These tags weighed 38 g in air with a length of 85 mm and a diameter of 16 mm. The tags transmitted information on tag depth (in depth bins with a resolution of 7 m) every 8 seconds from 7:00 a.m. to 7:00 p.m. AKDT. Tags were programmed to transmit on a 20 week on/off period beginning 1 July 2013.

2.3.6 Acoustic tracking

Acoustic tracking was conducted using an 8.5 m aluminum vessel. Two hydrophone configurations were employed that provided either fine-scale position estimates for tagged fish at slow vessel speeds (maximum 2.5 knots) or coarser-scale positions along transects at faster vessel speeds (maximum 8 knots). For the fine-scale configuration, two Lotek LHP1 hydrophones were deployed on each side of the vessel from an aluminum outrigger mounted on the bow. Each hydrophone was attached to a weight-bearing line 1 m above a 20-lb lead cannonball. Hydrophones were approximately 3 m apart at a depth of 3 m when the vessel was under way. This hydrophone configuration allowed real-time navigation toward tagged fish using a Lotek MAP 600 RT two-port receiver. For conducting higher-speed (5 – 8 knots) searches along transects, one hydrophone was attached to a weight-bearing line 1 m above a v-fin. The v-fin was towed 13 m behind the vessel at a depth of 12 m. Stationary test tags at 2 depths (approximately 100 m and 50 m) were deployed on a temporary fixed mooring during each trip to determine detection range, the relationship between signal strength and distance from the test tags, and

the accuracy of real-time navigation. Uncertainty in fine-scale position estimates was determined primarily by triangulation of angles to the tagged animal from different vessel trajectories while uncertainty in position estimates for high-speed tracking was determined primarily by signal strength. Additional information on acoustic tracking methods and detection distances is provided in Supplement S2-1.

Four tracking trips were conducted at 2-week intervals following the release of tagged animals in 2013, and two trips were conducted in 2014 during June and early July. During each 2 – 3-day tracking trip, the core study area (where tagged fish were released) was searched with both fine-scale and course-scale positioning methods. First, the core study area was systematically searched with high-speed tracking along linear transects spaced 400 m apart to obtain course-scale position estimates. Prior to initiating transects, a vessel speed was chosen that resulted in a detection distance of at least 200 m based on detection of stationary test tags. Low speed (fine-scale) tracking with the dual bow-mounted hydrophones was then conducted in the vicinity of each animal that had been located in the course-scale transects. After searching the core study area, adjacent areas at similar depths and habitat types were searched with high-speed tracking in an attempt to locate animals that were not detected in the core study area. Approximately 1/3 of the tracking time was spent searching areas outside of the core study area.

2.3.7 Study area depth and temperature data

A long-term stationary mooring containing four Desert Star PSATs was deployed in the core study area (Figure 2.1C) from October 7, 2013 to July 2, 2014 to serve as a "no movement" control and to allow comparisons of PSAT data to known depth and temperature data inside the MPA. PSATs were configured with 4-minute measurement intervals (the same as the tagged fish). Only depth and temperature data from these tags were used for data analysis. The sub-surface mooring consisted of an anchor, an acoustic release on a stainless steel swivel, a test tag line, and three trawl floats. The mooring was deployed at a depth of 147 m. Tags were firmly attached to the test tag line in a vertical orientation at depths from 130 – 134 m.

Oceanographic surveys have been conducted by Glacier Bay National Park personnel at 22 stations (including two stations outside the MPA in the adjacent Icy Strait) since 1992 (Johnson and Sharman, 2014). At each survey station, a CTD (Conductivity, Temperature, Depth) instrument is lowered through the water column, providing a vertical profile of temperature at every meter of depth sampled. Eight oceanographic survey stations, including Station 4 in the core study area (Figure 2.1B), are sampled every month from March to October and once in mid-winter. All 22 stations are sampled in mid-winter and mid-summer. Data for all years and stations from 1992 to 2014 were downloaded from the NPS website at http://science.nature.nps.gov/im/units/sean/OC_main.aspx.

To allow comparisons of PSAT data to temperature-at-depth data from locations outside of the MPA, long-term stationary mooring data were obtained at two locations in the Gulf of Alaska during the time period of the study (B. Stone, NOAA AFSC, Juneau, pers. comm.). Moorings were located on the Fairweather Grounds (58.2 N, 138.9 W) at a depth of 157 m and at Shutter Ridge (56.2 N, 135.1 W) at a depth of 208 m (Figure 1.1A). Star Oddi DST centi data loggers were deployed on each mooring to record temperatures (resolution 0.032°C, accuracy ± 0.1°C) at six-hour intervals.

2.3.8 Data analysis

Annual movement patterns at multiple scales were analyzed based on 1) within-summer movement patterns from acoustic tracking during 2013, 2) distance moved from the summer 2013 release location to end locations in winter 2014 and summer 2014, and 3) determining whether tagged fish with PSAT data left the MPA during the winter on spawning migrations by inferring whether the fish were inside or outside of the MPA for each day of their tag records.

2.3.8.1 Within-summer movement patterns

Movement patterns of acoustic-tagged fish during summer were characterized as either "home range" or "dispersive" based on previously-characterized fine-scale movement behavior of adult halibut within the MPA (Nielsen et al., 2014). Net Squared Displacement (NSD) analyses can be used to characterize and quantify movement patterns by fitting models that describe the relationship between the squared distance from the initial location to each subsequent relocation of the tagged animal over time

(Börger and Fryxell, 2012). For a fish that exhibits home range behavior, NSD does not increase over time so the spatial scale of home range behavior is estimated by fitting an intercept-only linear model to NSD vs. time. We used a mixed effects model to fit an intercept-only linear model to log(NSD) vs. time (with fish ID as a grouping variable) to determine the population average home range size (fixed effect) and estimate the size of each individual's home range (random effect). Observations obtained during the first three days following tagging were discarded to minimize the potential influence of temporary tagging effects on fish behavior in movement scale analyses. No attempt was made to quantify scales of dispersive movement, as observations were not frequent enough to determine whether dispersive movement was random or directed (Nielsen et al., 2014).

2.3.8.2 Seasonal and annual displacement

To quantify horizontal displacement from the release location from summer to winter and from summer to the following summer, we calculated horizontal displacement distances from release locations for each season using acoustic tracking, PSAT pop-up, and fisheries recapture locations. Summer-to-winter displacement was calculated based on PSAT pop-up or fisheries recapture locations from January 30 – March 9, 2014. Summer-to-summer displacement was calculated using the most precise observation available from acoustic telemetry, PSAT pop-up locations, or fisheries recaptures from May 20 – July 8, 2014. We calculated the median, minimum, and maximum horizontal displacement values observed for each season for both residents and migrants.

2.3.8.3 MPA residency and migration

Known locations (from telemetry, PSAT pop-ups, and fisheries recaptures) and PSAT depth and temperature data were used to determine whether or not tagged fish left the MPA during the winter, and if so, to estimate dates for exit and re-entry. To accomplish this, we developed a procedure to infer whether fish with PSAT data were likely to be either inside or outside of the MPA for each day of their archival tag record based on criteria derived from Glacier Bay's unique and well-characterized glacial oceanography and bathymetry (details provided in Supplement S2-2). Specifically, when the fish resided within the MPA, temperature-depth profiles (TDPs) from tagged fish matched unique temperature-depth

profiles measured within the MPA obtained from regular oceanographic sampling (Danielson, 2012; Johnson and Sharman, 2014) and the stationary mooring data (Figure 2.2). In addition, for fish with detailed data (available from physically recovered tags), tidal amplitude and phase in the depth record of resident fish matched the tide amplitude and phase in the depth records of the stationary mooring (Figure 2.3A). When fish left the MPA, TDPs and tidal amplitude/phase from PSAT records did not match known temperature and depth records within the MPA (Figures 2.2A and 2.2C), and tidal amplitude/phase did not match the stationary mooring (Figure 2.3B). In addition, tagged fish that recorded temperatures > 5.5°C at depths > 100 m in January or February could not have resided within the MPA during that time period based on a temperature threshold for winter residency derived from an analysis of historical oceanographic survey data within the MPA (Supplement S2-3). When a tagged fish crossed the shallow (40 m) glacial sill at the entrance to the MPA (Figure 2.1B), a maximum daily depth less than 70 m (a conservative estimate of sill crossing that accounts for uncertainties in depth accuracy of the tag, tidal stage, and accuracy of the bathymetric grid) was recorded in its PSAT depth record. Therefore, dates of exit and re-entry could be estimated based on dates when the maximum daily depth was less than 70 m followed by diverging depth and temperature records recorded by PSATs from TDP data collected within the MPA following the exit, or converged with TDP data from the MPA following re-entry (Figures 2.2A and 2.2C). Using this procedure, each fish with PSAT data was classified as either a resident or a migrant. We calculated the proportion of resident and migrant fish and compared average size (total length) of fish assigned to each behavioral type using a Welch 2-sample t-test after checking assumptions of normality and equal variance. In addition, we aggregated data from all fish for which archival records were available to provide a total number of fish-days during which fish were 1) inside the MPA, 2) outside the MPA while the commercial fishery was closed, and 3) outside the MPA when the commercial fishery was open and thus vulnerable to commercial fishing.

We looked for spawning activity in the form of spawning rises in detailed data sets of both resident and migrant fish. Spawning rises consisted of an evenly spaced series of abrupt rises into the water column at approximately three-day intervals thought to represent at least one type of reproductive

behavior that can be identified in archival tag records (Seitz et al., 2005). Timing and depth of individual rises and total time between the first and final rises were recorded.

We summarized trends in depth and temperature experienced by both resident and migrant fish by creating composite TDPs for each behavioral type. Composite TDPs were created by averaging the maximum temperature recorded by individual fish in each 5-m depth bin for each day of the study. The resident TDP composite provided information on seasonal patterns in the water column such as warming and stratification of the water column throughout the summer and into the fall, followed by mixing and cooling of the water column late in the year. To compare resident and migrant TDPs to known conditions within the MPA and in the Gulf of Alaska, stationary mooring data from the core study area and two stations in the Gulf of Alaska were overlaid on each composite TDP (details provided in Supplement S2-2). Plots of standard deviation in maximum temperature for each 5-m depth bin were also produced (when more than one observation was available in each depth bin) to provide information on the variation in composite TDPs. The overall difference in temperatures observed for each type of behavior was determined by subtracting the resident composite from the migrant composite. Dates of exit/re-entry for migrating fish were superimposed on the plot to visualize migration timing relative to seasonal changes in water column temperatures.

2.4 Results

2.4.1 Fish capture and tagging

The average total length for tagged fish was 131 cm (s.d. 17 cm) and ranged from 105 to 170 cm (Table 2.1). The average time on board the vessel during the tagging procedure was two minutes for those fish that received only PSATs and seven minutes for those that received both PSATs and acoustic tags. The average depth at release locations was 41 m (s.d. 12 m).

PSAT depth and temperature data were received for 21/25 tagged fish (WC = 10/10, DS = 11/15). Thirteen PSATs were physically recovered (WC = 5, DS = 8) and thus provided detailed data sets (10-second and 4-minute recording intervals, respectively). End locations were available for 9/10 WC tags

and 7/15 DS tags (Table 1.1). On five occasions, PSAT data were obtained but end locations were not available (Fish # 4, 17, 18, 19, and 25).

2.4.2 Within-summer movement patterns

In general, tagged fish could be detected acoustically if they were within 200 to 500 m of the tracking vessel. However, detection distance varied with hydrophone configuration (v-fin vs. fine-scale), speed of vessel travel, environmental conditions, and the depth of the tagged fish (Supplement S2-1). The average estimated uncertainty for positions obtained with the fine-scale tracking was 76 m (s.d. 39 m). The average estimated uncertainty of v-fin (course-scale) estimates was 143 m (s.d. 37 m).

Almost all (14/15) of the acoustic-tagged fish were detected at least once during the 2-month acoustic tracking period in 2013. Non-dispersive, home range behavior in the release location was observed for 12/15 acoustic-tagged fish (Figure 2.4A). The population average home range size (fixed effect) estimate from the NSD mixed effects model was 12.2 m^2 (s.e. 0.5) on the log scale (Figure 2.4B), which corresponds to an average movement radius of 485 m (95% C.I. 292 – 809 m). Estimated movement radii for individual fish (random effect) ranged from 86 to 1705 m. At least two animals (#24 and #16) exhibited dispersive movement and were found in locations outside of the core study area, but within the MPA, 7.8 and 11.8 km, respectively, from their release locations at the end of the active tracking season.

2.4.3 Seasonal and annual displacement

All of the acoustic, PSAT, and fisheries recapture locations used to calculate summer-to-winter and summer-to-summer displacement from the release location except one (the recapture location for the migratory Fish #23) were inside the MPA (Figure 2.5A). The median distance between the release location and winter PSAT pop-up or fisheries recapture locations was 4.7 km (range 1.4 – 48.1 km, n = 11). Three of the summer-to-winter locations used to calculate displacement were from fish that had undergone winter spawning migrations to locations outside of the MPA (Figure 2.5B). The median distance moved over the course of one year (summer-to-summer) was smaller than the movement between summer-to-winter (median 1.3 km, range 0.3 – 26.5 km, n = 14). Two of the fish in the summer-

to-summer treatment had undertaken migrations to locations outside of the MPA during the winter. However, the larger displacement values observed during the winter season may reflect the relative lack of precision from PSAT locations (maximum error of 5000 m for DS tags) compared to acoustic telemetry (maximum error of 200 m) that dominated the summer-to-summer displacement measurements (Table 2.2).

2.4.4 MPA residency and seasonal migration

The proportion of migratory fish was determined to be 0.29 (95% C.I. 0.09 – 0.48), as only six of the 21 fish with PSAT data were found to have conducted winter migration (Figure 2.6). Dates of departure from the MPA ranged from December 8 to December 30, 2013. Four of the six fish that departed the MPA in December returned by mid-March (range January 28 to March 15). Three of the migrants returned to the MPA within several days of each other (January 28–30). One fish (#23) was captured outside the MPA on March 9, and the fate of the last fish (#18) was unknown. Based on the dates of exit and re-entry of these four fish, the average time spent outside of the MPA was 57 days (s.d. 22 days).

Migrant fish were significantly larger than resident fish ($p = 0.038$). The average length of resident fish was 123.9 cm (s.d. 12.9, range 105 – 151, $n = 15$) compared to an average length of 138.2 cm (s.d. 12.2, range 119 – 152, $n = 6$) for migrants. The difference in size distribution between resident and migratory fish was driven by a larger proportion of smaller fish among the resident fish. When the four resident fish smaller than the smallest size of migratory fish (119 cm) were removed, there was no significant difference between the size of residential and migratory fish ($p = 0.15$).

Putative spawning behavior was observed for some migratory fish, but not for any MPA winter residents. Spawning rises were observed in 2/4 of detailed depth records available for migratory fish (#14 and #24). Spawning periods for both fish lasted 12 days, beginning on 17 and 9 January respectively, and consisted of five separate spawning events for each fish.

TDP composite summaries for resident (Figure 2.7A) and migrant (Figure 2.7B) fish were similar during the summer and fall, but diverged in December when migratory fish departed the MPA. Both

resident and migrant TDP composites revealed a trend of gradual warming of the water column between July and November. Individual variation in TDPs for both resident and migratory fish during this time period was small (Figures 2.7C and 2.7D). In December, when resident TDP composites indicated that the water column had become well-mixed, TDP composites for migrant fish began to diverge from the resident fish composites as migrant fish departed the MPA (Figure 2.7E). Migrant fish experienced water temperatures up to 2°C warmer than resident fish in January. In contrast to resident composite TDPs, which matched the temperature data from the stationary mooring in the core study area year-round, TDP data from migrating fish more closely resembled stationary mooring data from the Gulf of Alaska. In addition to differences in TDPs, tidal amplitudes observed for migrating fish were approximately half that of tidal amplitudes observed at the stationary mooring in the core study area (Figure 2.3B), which is similar to tidal amplitudes predicted in the Gulf of Alaska compared to the MPA (Shi et al., 2014).

Most tagged fish were not vulnerable to commercial halibut fishing, even during winter spawning migrations to locations outside of the MPA. Of the total 4690 fish-days from the aggregated archival records for 21 fish, 4354 (92.8%) were inside the MPA and 328 (7.0%) were outside the MPA when the commercial fishery was closed from November 7 – March 8. Only two fish were outside the MPA for a total of 8 fish-days (0.2%) while the commercial fishery was open. Thus, spawning migrations were conducted almost entirely within the period of time when the commercial fishery was closed.

2.5 Discussion

2.5.1 Spatial and temporal scales of site fidelity

This study corroborates the findings of summer home range behavior and other non-random movement patterns that were observed for similarly sized halibut within the MPA in the early 1990s (Nielsen et al., 2014). The proportion of fish with summer home range behavior was similar between the two studies (75%, 12/15, in this study vs. 63%, 27/43, in the previous study). The average home range radius in this study (485 m) was within the confidence intervals of the average home range radius found in the previous study (Figure 2.4B). The distances moved during dispersive movements were also similar between the two studies (a maximum displacement of 26.5 km for non-migratory fish compared to 18 km

in the previous study). This suggests that the proportion of fish with home range vs. dispersive movement patterns and the spatial scale of the movement patterns are relatively stable over time.

Through the use of both acoustic tracking and PSAT data, we were able to determine that the majority of tagged fish in the current study exhibited interannual site fidelity at spatial scales < 5 km even if they undertook migrations to locations outside of the MPA during the winter. This level of resolution was achieved through the use of complementary tagging approaches. Acoustic telemetry provided precise information on home range locations and spatial scales of movement during summer, while PSATs provided information on winter migrations. For example, Fish #14 was observed to display home range behavior in the study area following tagging with acoustic telemetry. PSAT data were used to detect departure from the MPA during December, spawning outside of the MPA in January, and a return to the MPA in March. In June of the following summer, Fish #14 was located within 1 km of its release location with acoustic telemetry. Previous PSAT studies in the Gulf of Alaska, Aleutian Islands, and the Bering Sea have documented fish undertaking seasonal spawning migrations followed by return to the summer feeding location (Loher, 2008; Loher and Blood, 2009; Seitz et al., 2011), but could not demonstrate the spatial and temporal scales of site fidelity with the level of detail available in the current study.

2.5.2 Migration characteristics

The results of this study support the existence of partial migration for halibut. Partial migration, where a proportion of individuals from the same population undertake migrations while others remain in localized areas year-round, is a common phenomenon for many fish species (Chapman et al., 2012). Previous studies found percentages of migrating halibut (of the same size class as this study) to be 84% of 76 fish tagged throughout the Gulf of Alaska (Loher and Seitz, 2008), 58% of 31 fish tagged from British Columbia to Kodiak, Alaska (Loher, 2011), and 85% of 48 fish tagged in the Bering Sea and Aleutian Islands (Seitz et al., 2011). One possible explanation for the smaller percentage of migrants observed in this study (29%, 95% C.I. 9% – 48%) is that the Glacier Bay MPA is a spatially constrained body of water, and therefore may have a higher proportion of year-round residency compared to more open and unconstrained habitats. For example, all (n = 4) Atlantic halibut tagged with PSATs in a Norwegian fjord

remained within the fjord throughout the spawning season and did not join known off-shore spawning aggregations (Seitz et al., 2014).

Another possible explanation for the small proportion of migrants is that some resident halibut may not have been sexually mature. Though the fish tagged in this study were likely to be mature based on data from the overall population in the Gulf of Alaska (Loher and Seitz, 2008), size at maturity may vary spatially (St-Pierre, 1984). Specifically, some inshore areas may contain increased proportions of larger fish that are not yet reproductively mature, relative to smaller average size-at-maturity offshore, and these inshore regions thus have lower migration rates compared to the overall Gulf of Alaska population (T. Loher, IPHC, Seattle, pers. comm.). The significantly larger body size of migrating adults compared to residents in this study supports this hypothesis. However, many of the fish that did not migrate were as large as the fish that did migrate, so this would not fully explain our findings of partial migration.

It is also possible that the partial migration observed in this study is caused by skip-spawning, which is thought to explain partial migration for breeding in other fish species (Chapman et al., 2012) and has been hypothesized for halibut (Loher and Seitz, 2008; Seitz et al., 2011). For those fish that skip spawning in a given year, undertaking a migration is likely an unnecessary energetic expense and may explain the observation of residents in this study. Because putative spawning behavior was observed for some migrating fish, the migrations observed in this study were likely to be associated with reproduction. However, it is not known whether the lack of spawning behavior observed in resident fish represents a lack of reproduction, or whether another form of spawning behavior exists that is not evident in archival tag data. A previous study on halibut seasonal spawning migrations found that serial spawning rises were observed in 10 of 14 detailed data sets for migrating fish (as defined by movement to depths greater than the continental shelf) but 0 of 2 detailed data sets for fish classified as residents (Loher and Seitz, 2008). Future studies with archival tags that measure high-resolution acceleration could help determine whether spawning behavior exists for resident fish or migrant fish that do not exhibit spawning rises.

Departure and return dates for fish migrating from Glacier Bay were within the range of departure and return dates observed in previous studies of migration timing for fish from the eastern and central Gulf of Alaska (Loher, 2011), where halibut migrations can begin as early as September and can end as late as April. However, the migrating fish from Glacier Bay exhibited a much narrower range of departure and return dates. This suggests that there could be a spatial component to migration timing, where fish from similar summer foraging areas have similar migration dates. Another possibility is that fish within the MPA are responding to distinct environmental cues for migration (e.g. strong gradients in oceanographic characteristics between the Glacier Bay MPA and the Gulf of Alaska), whereas fish in areas with weaker gradients may need to respond to broader cues for migration and return.

The six migratory fish all left in December, in close proximity to both a change in photoperiod (winter solstice) and seasonal mixing of the water column (approximately December 15; Figure 2.7A and 2.7B). Changes in photoperiod and water column temperature have been observed to be related to the onset of migration for other flatfish species (Gibson, 1997), so it is possible that one or both of these serves as a cue for winter migration initiation in halibut. Water temperature has been linked to migration timing of flounder, *Platichthys flesus* (L.) (Sims et al., 2004). There was more variation in the return date, with most returning at the end of January and one fish returning six weeks later. This variation might result from travel to different spawning locations or from different cues for the return portion of the migration. Further research to determine cues for migration will be important for understanding how the timing of migration may vary over space and between years.

Migrating fish likely visited locations in the Gulf of Alaska based on similarity of migrant TDPs to winter water temperatures at two locations in the Gulf of Alaska and differences in tidal amplitude/phase compared to the stationary mooring within the MPA. Most of the known spawning locations for halibut in Southeast Alaska are in the Gulf of Alaska along the edge of the continental shelf at depths greater than 180 m (St-Pierre, 1984). Some of the largest aggregations of spawning halibut have been observed in the Gulf of Alaska 100 – 300 km north of the Glacier Bay MPA. It is possible that some of the tagged fish from this study may have joined those aggregations.

2.5.3 Potential utility of protected areas for halibut

Our results on spatial and temporal scales of site fidelity, migration timing, and the proportion of migratory fish have important implications for the potential effectiveness of the Glacier Bay MPA for conservation of halibut spawning stock. In addition to confirming a tendency for limited dispersal and home range behavior during summer, this study found that fish with PSAT data spent most of their time (92.8% of all fish-days) within the boundaries of the MPA. The small proportion of tagged fish that left the MPA on spawning migrations did so while the commercial fishery was closed, and half of the migrating fish returned while the commercial fishery was still closed. In addition, most round-trip migrations largely coincided with the area-wide sport fishing closure during the month of January.

The Glacier Bay MPA could serve as a year-round refuge from commercial fishing for adult halibut after the phase-out of commercial fishing is complete (estimated to occur sometime after 2050, C. Soiseth, NPS, pers. comm.) due to the following conditions: 1) halibut exhibit behavior such as site fidelity and homing capabilities to locations within the MPA, 2) the spatial extent of the MPA encompasses the movement of halibut during the summer when the commercial fishery is open, and 3) temporal fisheries closures protect most migrating fish from harvest during the winter. Thus, potential effectiveness of protected areas for halibut depends on both the spatial extent of the protected area relative to scales of fish movement during the summer and appropriate timing of the large-scale commercial fishery closure during the winter. Because the commercial fishery closure period varies from year to year, and the timing of migrations may vary from year to year, the effectiveness of protected areas for halibut would also be expected to vary over time. If the winter commercial fishery closure is ever discontinued in response to increasing market demands for fresh halibut year-round (Loher, 2011), the effectiveness of protected areas would depend more strongly on the proportion of fish that undertake winter spawning migrations outside of protected areas.

MPAs may provide benefits such as increased size structure and abundance of fish that reside within their boundaries (Edgar et al., 2014). One important consideration in understanding whether such benefits could occur for halibut in the Glacier Bay MPA is that recreational fishing for halibut will

continue within the MPA after the commercial fishery ceases. Despite the lack of total protection from all fishery sectors, some benefits to halibut still could occur in this "partially protected" marine area due to reduced harvest within its boundaries compared to adjacent areas (Sciberras et al., 2015). Additional information on the magnitude of recreational harvest combined with information on abundance and size structure is needed to further understand overall effectiveness of the Glacier Bay MPA for halibut.

In addition to potential changes in abundance or size structure within the MPA, harvest refuges such as Glacier Bay may provide potential benefits in terms of enhancement of the range-wide spawning stock for halibut. Specifically, fish that leave the protected area to spawn are contributing to the larger population of halibut in the Gulf of Alaska. This could be viewed as an active form of "spillover", where eggs and larvae are exported from the MPA by adults that participate in spawning migrations. Therefore, if adult female fish are allowed to reach larger size, these fish will supply more eggs and larvae to the overall population. Because large females provide more eggs of higher quality, they may serve as a buffer for environmental variability (Hixon et al., 2014). As the halibut fishery is currently estimated to be comprised of 2/3 adult females and has recently experienced steep declines in spawning stock abundance (Stewart et al., 2016), spatial refuges that protect large females may provide opportunities for conservation of spawning stock.

2.5.4 Experimental design and caveats

Our study design offered several important advantages for detecting and characterizing seasonal migrations of adult halibut. First, fish were tagged within a deep (450 m) fjord with a relatively shallow (40 m maximum) sill at the entrance. This bathymetric feature provided a unique metric for identifying dates of MPA exit and re-entry. Second, TDPs and tidal amplitude/phase were well-characterized and distinct from adjacent areas due to the glacial influence and topographical features of the glacial fjord system. Third, the long-term stationary moorings in the release area and in the Gulf of Alaska served as effective tools for inferring residence within and outside of the MPA through direct comparison with tagged animal depth and temperature records. Finally, double-tagging fish with acoustic tags helped to increase the number of known locations as well as provide insights into behavioral patterns during the

summer. These advantages allowed us to obtain detailed information on migration timing and the proportion of halibut that participate in seasonal migrations, which is important for determining seasonal and spatial vulnerability of adult halibut to harvest (Loher, 2011).

Despite these experimental advantages, a few caveats exist. First, we acknowledge that fish thought to reside in the MPA until the February 1 PSAT release date could have left the MPA on spawning migrations after that date. However, in an archival tag study of halibut migration and spawn timing in the Gulf of Alaska (Loher, 2011), none of the 61 archival-tagged halibut that undertook spawning migrations began their migrations after January 15. Therefore, we believe it is likely that the halibut that remained in the MPA through February 1 did not leave the MPA later in the winter. Future studies should deploy more PSATs programmed to collect data for a full year to confirm this assumption. Second, the number of PSATs deployed in this study was enough to demonstrate that some fish migrate and others remain in the MPA during the winter, but the proportion of residents and migrants reported here is based on a relatively small sample size; thus, future studies with higher sample sizes would quantify the proportion of residents and migrants more precisely. Third, it is possible that the sizes of home ranges reported here are influenced by the size of the core area; i.e., if a fish left the core area as part of its home range behavior it would not be detected during tracking and thus home ranges may be biased low. However, the locations of most tagged fish were well within the spatial extent of tracked area, the home range sizes observed were consistent with those from the previous study (where tracking was not limited to a core study area), and annual (summer-to-summer) displacement distances which included PSAT pop-up and fisheries recapture locations occurred at spatial scales similar to home ranges determined by acoustic tracking (e.g., < 2 km). Fourth, because the fish tagged in this study were large (> 105 cm) and likely to be adult female, spatial and temporal scales of movement may differ for smaller fish such as males or juveniles. Because the commercial fishery targets halibut > 82 cm, the results provided here may not apply to all fish vulnerable to the commercial fishery. Finally, the duration of the study (one year) may provide an incomplete picture of the movement patterns for individual fish, as movement patterns could change over time scales of multiple years (for example, migration could occur

every other year for skip-spawning fish). Nevertheless, this study provides important observations of site fidelity and homing in a migratory species that will benefit fisheries management and lead to future research on site fidelity behavior in the overall population.

2.6 Supplementary material

Three appendices are available as supplementary material at ICESJMS online. Supplement S2-1 provides additional details about acoustic tracking methods and detection distances. Supplement S2-2 describes the procedure for inferring residency and migration based on PSAT depth and temperature data. Supplement S2-3 provides a historical analysis of National Park Service oceanographic survey data used to derive the maximum temperature criterion for winter residence within the MPA.

2.7 Acknowledgements

We would like to thank the National Park Service, the Pollock Conservation Cooperative Research Center, the Rasmuson Fisheries Research Center, the North Pacific Research Board, the UAF Undergraduate Research and Scholarly Activity program, and the National Undersea Research Program for their support of this project. We would also like to thank Zach Stenson (F/V Taurus), Mark Evans, Kevin Siwicke, Michael Courtney, Anne Beaudreau, Gordon Kruse, Jessica Glass, John Rodstrom, Amanda Compton, Thomas Farrugia, Maurice Tivey, David Stone, Barbara Stone, Alan Steffert, Tim Loher, Craig Murdoch, Chad Soiseth, Julien Appignanii, Katy Rayfield, Jim de La Bruere and the crew of R/V Medeia for help with the fieldwork. We thank Milo Adkison, Franz Mueter, Susanne McDermott, Tim Loher, Chad Soiseth, and Craig Murdoch for advice and feedback on the project design and manuscript.

2.8 References

Börger, L., and Fryxell, J. 2012. Quantifying individual differences in dispersal using net squared displacement. *In* Dispersal and spatial evolutionary ecology, pp. 222 - 230. Ed. by J. Clobert, M. Baguette, T. Benton, and J. Bullock. Oxford University Press.

Chapman, B. B., Skov, C., Hulthen, K., Brodersen, J., Nilsson, P. A., Hansson, L. A., and Bronmark, C. 2012. Partial migration in fishes: definitions, methodologies and taxonomic distribution. Journal of Fish Biology, 81: 479-499.

Clark, W. G., and Hare, S. R. 2006. Assessment and management of Pacific halibut: data, methods, and policy. Scientific Report No. 83, International Pacific Halibut Commission, Seattle, WA. 104 pp.

Danielson, S. L. 2012. Glacier Bay oceanographic monitoring program analysis of observations, 1993–2009. Natural Resource Technical Report NPS/SEAN/NRTR—2012/527. National Park Service, Fort Collins, Colorado. 172 pp.

Edgar, G. J., Stuart-Smith, R. D., Willis, T. J., Kininmonth, S., Baker, S. C., Banks, S., Barrett, N. S., et al. 2014. Global conservation outcomes depend on marine protected areas with five key features. Nature, 506: 216-220.

Evans, K., and Arnold, G. 2009. Summary Report of a workshop on geolocation methods for marine animals. *In* Tagging and Tracking of Marine Animals with Electronic Devices, pp. 343-363. Ed. by J. L. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert. Springer Netherlands, Dordrecht.

Gibson, R. N. 1997. Behaviour and the distribution of flatfishes. Journal of Sea Research, 37: 241-256.

Grüss, A., Kaplan, D. M., and Hart, D. R. 2011. Relative impacts of adult movement, larval dispersal and harvester movement on the effectiveness of reserve networks. PLoS ONE, 6: e19960.

Halpern, B. S. 2003. The impact of marine reserves: Do reserves work and does reserve size matter? Ecological Applications, 13: S117-S137.

Hixon, M. A., Johnson, D. W., and Sogard, S. M. 2014. BOFFFFs: on the importance of conserving old-growth age structure in fishery populations. ICES Journal of Marine Science, 71: 2171-2185.

Johnson, W. F., and Sharman, L. C. 2014. Glacier Bay National Park and Preserve oceanographic monitoring protocol: Version OC-2014.1. Natural Resource Report NPS/SEAN/NRR — 2014/851, National Park Service, Fort Collins, Colorado. 259 pp.

Lascelles, B., Notarbartolo Di Sciara, G., Agardy, T., Cuttelod, A., Eckert, S., Glowka, L., Hoyt, E., et al. 2014. Migratory marine species: their status, threats and conservation management needs. Aquatic Conservation: Marine and Freshwater Ecosystems, 24: 111-127.

Loher, T. 2008. Homing and summer feeding site fidelity of Pacific halibut (*Hippoglossus stenolepis*) in the Gulf of Alaska, established using satellite-transmitting archival tags. Fisheries Research, 92: 63-69.

Loher, T. 2011. Analysis of match–mismatch between commercial fishing periods and spawning ecology of Pacific halibut (*Hippoglossus stenolepis*), based on winter surveys and behavioural data from electronic archival tags. ICES Journal of Marine Science, 68: 2240-2251.

Loher, T., and Blood, C. 2009. Dispersion of Pacific halibut (*Hippoglossus stenolepis*) summering off British Columbia and the US Pacific Northwest, evaluated via satellite archival tagging. Canadian Journal of Fisheries and Aquatic Sciences, 66: 1409-1422.

Loher, T., and Seitz, A. C. 2006. Seasonal migration and environmental conditions of Pacific halibut *Hippoglossus stenolepis*, elucidated from pop-up archival transmitting (PAT) tags. Marine Ecology Progress Series, 317: 259-271.

Loher, T., and Seitz, A. C. 2008. Characterization of active spawning season and depth for eastern Pacific halibut (*Hippoglossus stenolepis*), and evidence of probable skipped spawning. Journal of Northwest Atlantic Fishery Science, 41: 23-36.

Nielsen, J. K., Hooge, P. N., Taggart, S. J., and Seitz, A. C. 2014. Characterizing Pacific halibut movement and habitat in a Marine Protected Area using net squared displacement analysis methods. Marine Ecology Progress Series, 517: 229-250.

Sciberras, M., Jenkins, S. R., Mant, R., Kaiser, M. J., Hawkins, S. J., and Pullin, A. S. 2015. Evaluating the relative conservation value of fully and partially protected marine areas. Fish and Fisheries, 16: 58-77.

Seitz, A. C., Loher, T., Norcross, B. L., and Nielsen, J. L. 2011. Dispersal and behavior of Pacific halibut *Hippoglossus stenolepis* in the Bering Sea and Aleutian Islands region. Aquatic Biology, 12: 225-239.

Seitz, A. C., Michalsen, K., Nielsen, J. L., and Evans, M. D. 2014. Evidence of fjord spawning by southern Norwegian Atlantic halibut (*Hippoglossus hippoglossus*). ICES Journal of Marine Science, 71: 1142-1147.

Seitz, A. C., Norcross, B. L., Wilson, D., and Nielsen, J. L. 2005. Identifying spawning behavior in Pacific halibut, *Hippoglossus stenolepis*, using electronic tags. Environmental Biology of Fishes, 73: 445-451.

Seitz, A. C., Wilson, D., Norcross, B. L., and Nielsen, J. L. 2003. Pop-up archival transmitting (PAT) tags: a method to investigate the migration and behavior of Pacific halibut *Hippoglossus stenolepis* in the Gulf of Alaska. Alaska Fishery Research Bulletin, 10: 124–136.

Shi, L., Wang, J., Myers, E., and Huang, L. 2014. Development and use of tide models in Alaska supporting VDatum and hydrographic surveying. Journal of Marine Science and Engineering, 2: 171-193.

Sims, D. W., Wearmouth, V. J., Genner, M. J., Southward, A. J., and Hawkins, S. J. 2004. Low-temperature-driven early spawning migration of a temperate marine fish. Journal of Animal Ecology, 73: 333-341.

Skud, B. E. 1977. Drift, migration, and intermingling of Pacific halibut stocks. Scientific Report No. 63, International Pacific Halibut Commission, Seattle, WA. 42 pp.

St-Pierre, G. 1984. Spawning locations and season for Pacific halibut. Scientific Report No. 70, International Pacific Halibut Commission, Seattle, WA. 46 pp.

Stewart, I. J., and Monnahan, C. C. 2016. Overview of data sources for the Pacific halibut stock assessment and related analyses. *In* Report of Assessment and Research Activities 2015, pp. 99 - 187. Ed. by L. Sadorus. International Pacific Halibut Commission, Seattle WA.

Stewart, I. J., Monnahan, C. C., and Martell, S. 2016. Assessment of the Pacific halibut stock at the end of 2015. *In* Report of Assessment and Research Activities 2015, pp. 188 - 209. Ed. by L. Sadorus. International Pacific Halibut Commission, Seattle WA.

Valero, J. L., and Webster, R. A. 2012. Current understanding of Pacific halibut migration patterns. *In* Report of Assessment and Research Activities 2011, pp. 341-379. Ed. by L. Sadorus. International Pacific Halibut Commission, Seattle, WA.

Webster, R. A., Clark, W. G., Leaman, B. M., and Forsberg, J. E. 2013. Pacific halibut on the move: a renewed understanding of adult migration from a coastwide tagging study. Canadian Journal of Fisheries and Aquatic Sciences, 70: 642-653.

Witherell, D., and Woodby, D. 2005. Application of marine protected areas for sustainable production and marine biodiversity off Alaska. Marine Fisheries Review, 67: 1-27.

Table 2.1. Information on fish tagged with PSATs (Pop-up Satellite Archival Tags), including: date of tagging, halibut size (total length), type of PSAT (WC = Wildlife Computers, DS = Desert Star), type of data (Detailed: high resolution data from physically recovered tag; Argos: daily summary data from satellites), programmed pop-up date, end location position (pop-up location), distance from the release location, and number of days tagged fish were at liberty.

Fish ID	Tagging Date	Fish T.L. (cm)	PSAT type	Data type	Assigned Pop-up date	End location date	End location lat	End location lon	Dist from release (km)	Days at liberty
1	6/29/13	137	WC	Argos	2/1/14	2/1/14	58.58650	-136.07720	4.6	217
2	6/29/13	106	WC	Detailed	7/1/14	7/1/14	58.75774	-136.52362	26.5	367
3	6/29/13	151	WC	Detailed	7/1/14	7/1/14	58.62159	-136.14005	0.3	367
4	6/29/13	127	WC	Detailed	7/1/14	N/A	N/A	N/A	N/A	328
5	6/30/13	132	WC	Argos	2/1/14	2/1/14	58.63450	-136.06000	4.7	216
6	6/30/13	152	WC	Argos	2/1/14	2/1/14	58.50600	-136.08410	12.8	216
7	6/30/13	105	WC	Detailed	2/1/14	2/1/14	58.62420	-136.12470	4.8	216
8	6/30/13	120	WC	Detailed	2/1/14	2/1/14	58.64620	-136.14730	2.5	216
9	6/30/13	124	WC	Argos	2/1/14	2/1/14	58.81140	-136.47800	24.3	216
10	6/30/13	116	WC	Argos	2/1/14	2/1/14	58.64820	-136.16580	1.5	216
11	7/1/13	127	DS	Argos	2/1/14	1/30/14	58.61556	-136.19222	3.3	213
12	7/1/13	170	DS	N/A	7/1/14	N/A	N/A	N/A	N/A	N/A
13	7/1/13	120	DS	Argos	2/1/14	2/1/2014 *	58.58750	-136.16850	6.8	215
14	7/1/13	139	DS	Detailed	7/1/14	7/2/2014 **	58.64970	-136.21222	1.5	366
15	7/1/13	128	DS	Detailed	7/1/14	7/1/14	58.70305	-136.07942	8.2	365
16	7/2/13	105	DS	N/A	7/1/14	N/A	N/A	N/A	N/A	N/A
17	7/2/13	114	DS	Detailed	2/1/14	N/A	N/A	N/A	N/A	N/A
18	7/2/13	132	DS	Detailed	2/1/14	N/A	N/A	N/A	N/A	N/A
19	7/2/13	120	DS	Detailed	2/1/14	N/A	N/A	N/A	N/A	214
20	7/2/13	151	DS	N/A	2/1/14	N/A	N/A	N/A	N/A	N/A
21	7/2/13	150	DS	N/A	2/1/14	N/A	N/A	N/A	N/A	N/A
22	7/2/13	148	DS	Detailed	2/1/14	2/1/2014 *	58.61950	-136.20340	4.7	214
23	7/2/13	119	DS	Detailed	7/1/14	3/9/14	58.28200	-136.19200	48.1	250
24	7/2/13	150	DS	Detailed	7/1/14	5/20/14	58.54633	-136.10133	10.8	322
25	7/2/13	130	DS	Argos	7/1/14	N/A	N/A	N/A	N/A	N/A
* 12-hour delay between first transmission and Argos location										
** 36-hour delay between first transmission and location of tag on shore										

Table 2.2. Seasonal and annual horizontal displacement of halibut from the Glacier Bay MPA, including: date of release and location, method of determining end location (PSAT = pop-up satellite end location determined by Argos, fishery = recapture location, acoustic = position estimate from acoustic tracking), days at liberty, horizontal displacement from release location (km), estimated error in location (m), and whether fish location was inside or outside of the MPA. Note that tag IDs 3 and 7 appear twice, first on their pop-up dates and second as recaptures.

Fish ID	Release date	End loc date	End loc type	Days at liberty	Dist moved (km)	Est Error (m)	Inside/ outside
Summer-to-winter displacement							
1	6/29/13	2/1/14	PSAT	217	4.6	300	Inside
5	6/30/13	2/1/14	PSAT	216	4.7	300	Inside
6	6/30/13	2/1/14	PSAT	216	12.8	300	Inside
7	6/30/13	2/1/14	PSAT	216	4.9	300	Inside
8	6/30/13	2/1/14	PSAT	216	2.5	500	Inside
9	6/30/13	2/1/14	PSAT	216	24.3	500	Inside
10	6/30/13	2/1/14	PSAT	216	1.5	500	Inside
11	7/1/13	1/30/14	PSAT	213	3.3	1500	Inside
13	7/1/13	2/1/14	PSAT	215	6.8	5000	Inside
22	7/2/13	2/1/14	PSAT	214	4.7	5000	Inside
23	7/2/13	3/9/14	Fishery	250	48.1	1000	Outside
Summer-to-summer displacement							
2	6/29/13	7/1/14	PSAT	367	26.5	500	Inside
3	6/29/13	7/1/14	PSAT	367	0.3	300	Inside
7	6/30/13	7/8/14	Fishery	373	4.0	1000	Inside
11	7/1/13	6/28/14	Acoustic	362	0.9	200	Inside
12	7/1/13	6/28/14	Acoustic	362	1.0	100	Inside
13	7/1/13	6/29/14	Acoustic	363	0.7	200	Inside
14	7/1/13	7/1/14	Acoustic	365	2.3	150	Inside
15	7/1/13	7/1/14	PSAT	365	8.2	2000	Inside
17	7/2/13	6/29/14	Acoustic	362	0.4	100	Inside
19	7/2/13	6/4/14	Acoustic	337	2.2	150	Inside
20	7/2/13	7/1/14	Acoustic	364	1.1	200	Inside
21	7/2/13	7/2/14	Acoustic	365	0.8	150	Inside
22	7/2/13	7/2/14	Acoustic	365	1.6	200	Inside
24	7/2/13	5/20/14	Fishery	322	10.8	1000	Inside
Summer-to-summer displacement (multi-year)							
3	6/29/13	6/5/15	Fishery	706	1.1	1000	Inside

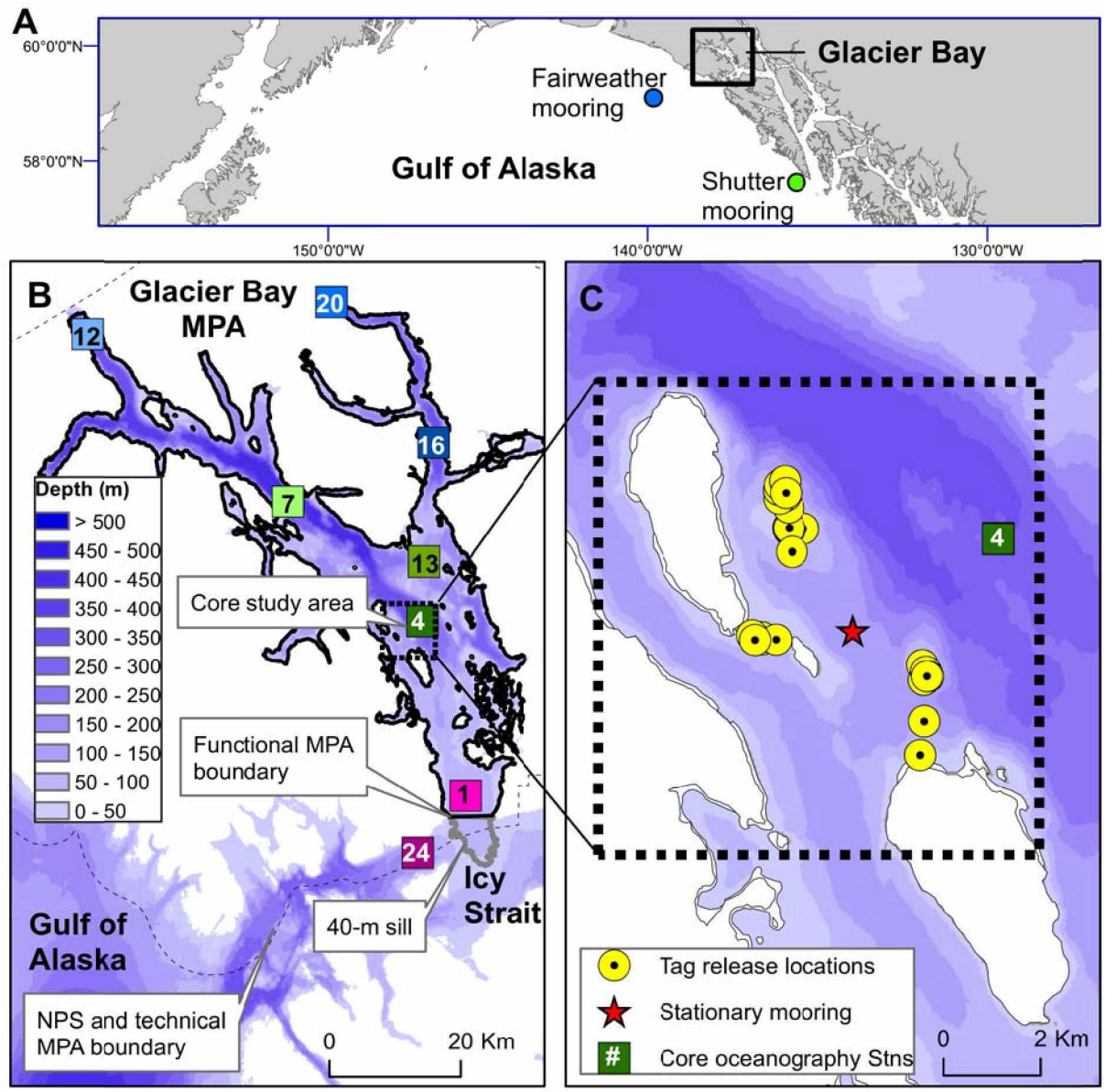


Figure 2.1. A) Location of Glacier Bay National Park and locations for two long-term stationary moorings (courtesy of Bob Stone, NOAA) in Gulf of Alaska. B) Glacier Bay functional MPA for halibut. Colored squares represent core National Park Service oceanographic monitoring stations. The entrance to the functional MPA is shown by a solid black line, and the 40-m sill that represents the boundary in the PSAT data is indicated by a solid gray line. C) The core study area (thick dashed line) within the MPA. Release locations for tagged fish are shown in yellow, the location of the long-term stationary mooring is shown by a red star, and NPS oceanography Station 4 where CTD casts are regularly conducted is shown by a green square.

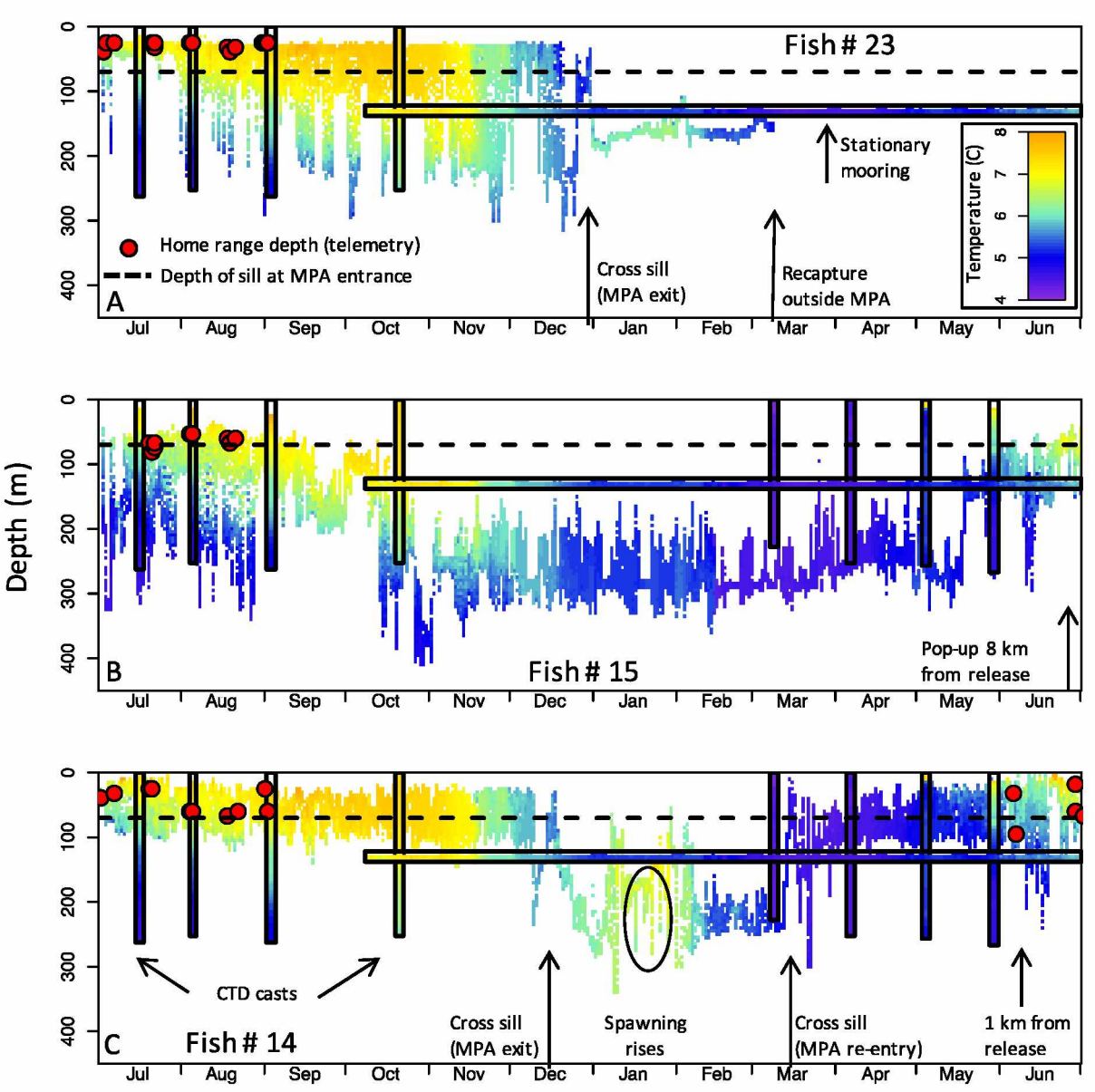


Figure 2.2. Example of matching temperature-depth profiles from tagged halibut to CTD casts (vertical bars) and stationary mooring data (horizontal bar) in the core study area to determine movement, including A) a migrant (Fish #23) with a known winter location outside of the MPA, B) a resident (Fish #15), for which fish TDPs matched TDPs in the core area year round, and C) a returning migrant (Fish #14), for which fish TDPs did not match the core area TDPs during the winter. Red circles indicate depths transmitted by acoustic tags during tracking trips; all three fish had home range behavior in the study area during 2013. Dashed line at 70 m indicates the criterion for sill crossings to determine dates for MPA exit and re-entry.

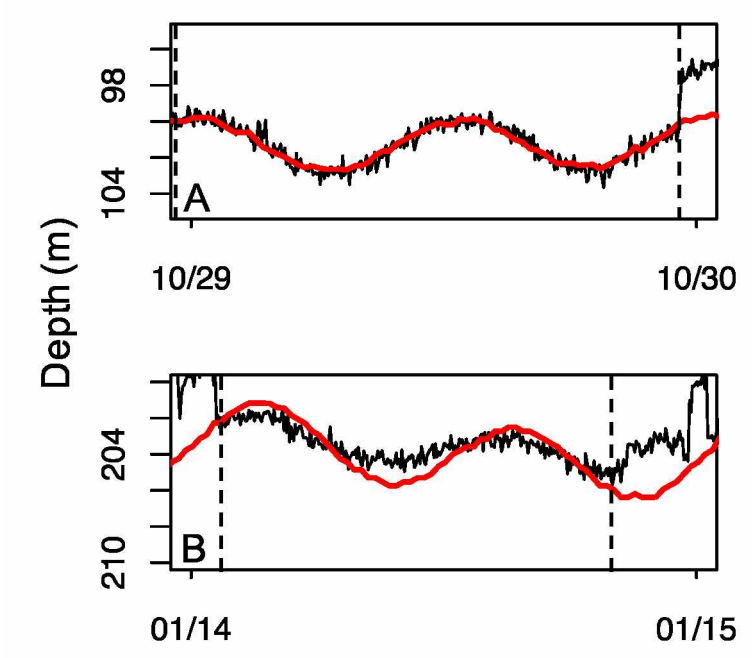


Figure 2.3. Comparison of tidal amplitude and phase in the depth records from the PSAT data of a migratory fish (Fish #24) with depth data from the stationary mooring in the core study area. In both plots, the fish was stationary for the time period between the two vertical dashed lines. A) Prior to migration, when the fish was inside the MPA, PSAT depth data (thin black line) matched changes in depth measured by the stationary mooring tag that was offset to match the depth of the PSAT (thick red line). B) In the time periods between spawning rises, when the fish was outside of the MPA, PSAT depth data did not match the tidal amplitude and phase recorded in the core study area.

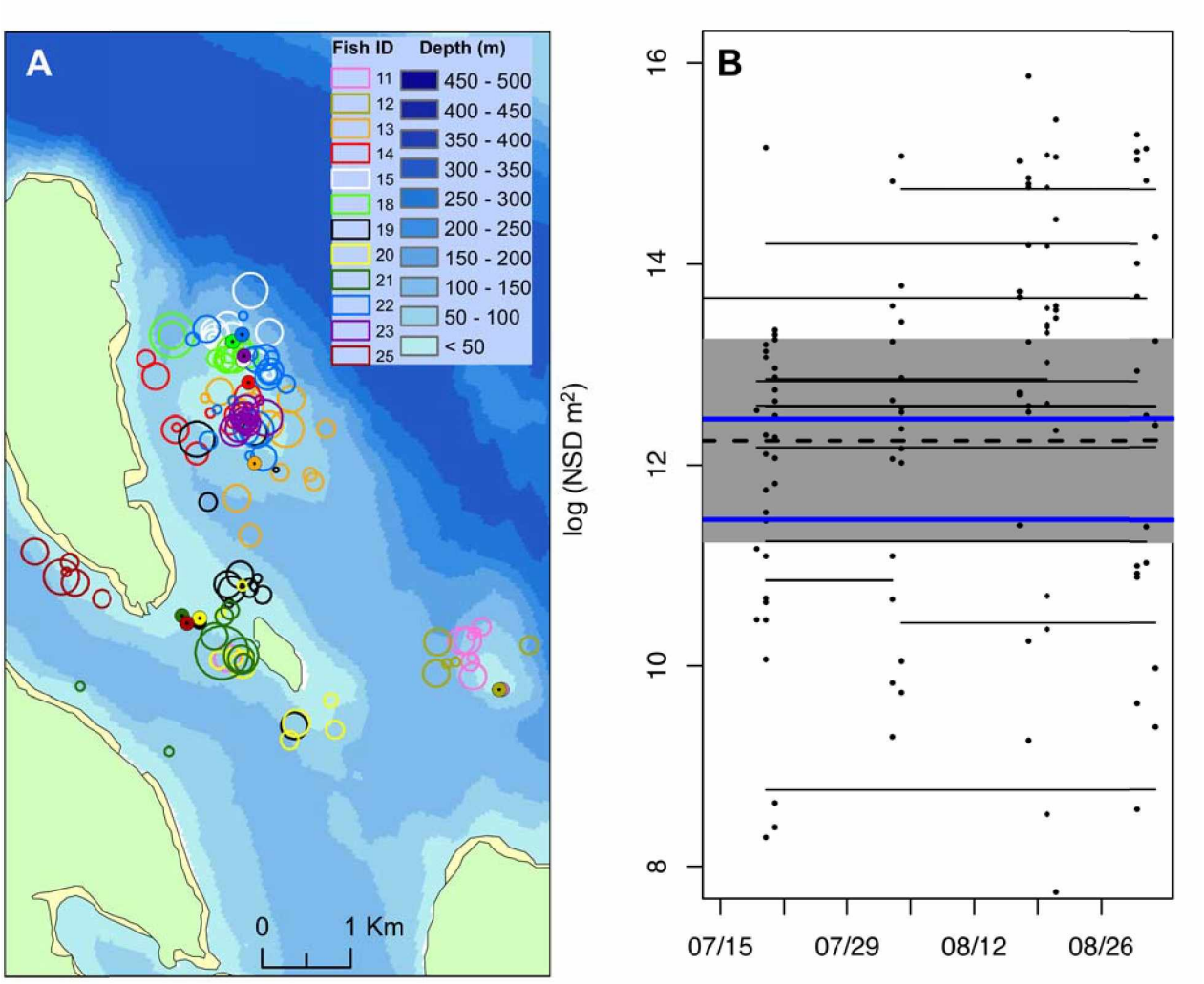


Figure 2.4. Home range behavior of acoustic-tagged halibut during summer 2013. A) Release locations (solid symbols) and estimated locations (hollow circles) of individual fish on all 2013 tracking trips combined. The size of the circle corresponds to the error in each position estimation. B) Home range size during summer 2013 determined by NSD analysis. Thick black dashed line represents population estimate for home range size (gray rectangle indicates 95% C.I.), and thin black lines represent individual home ranges. Points indicate observations of individual fish. Thick blue lines indicate 95% C.I.s for the average home range size observed during previous research in the study area (Nielsen et al., 2014, Chapter 1).

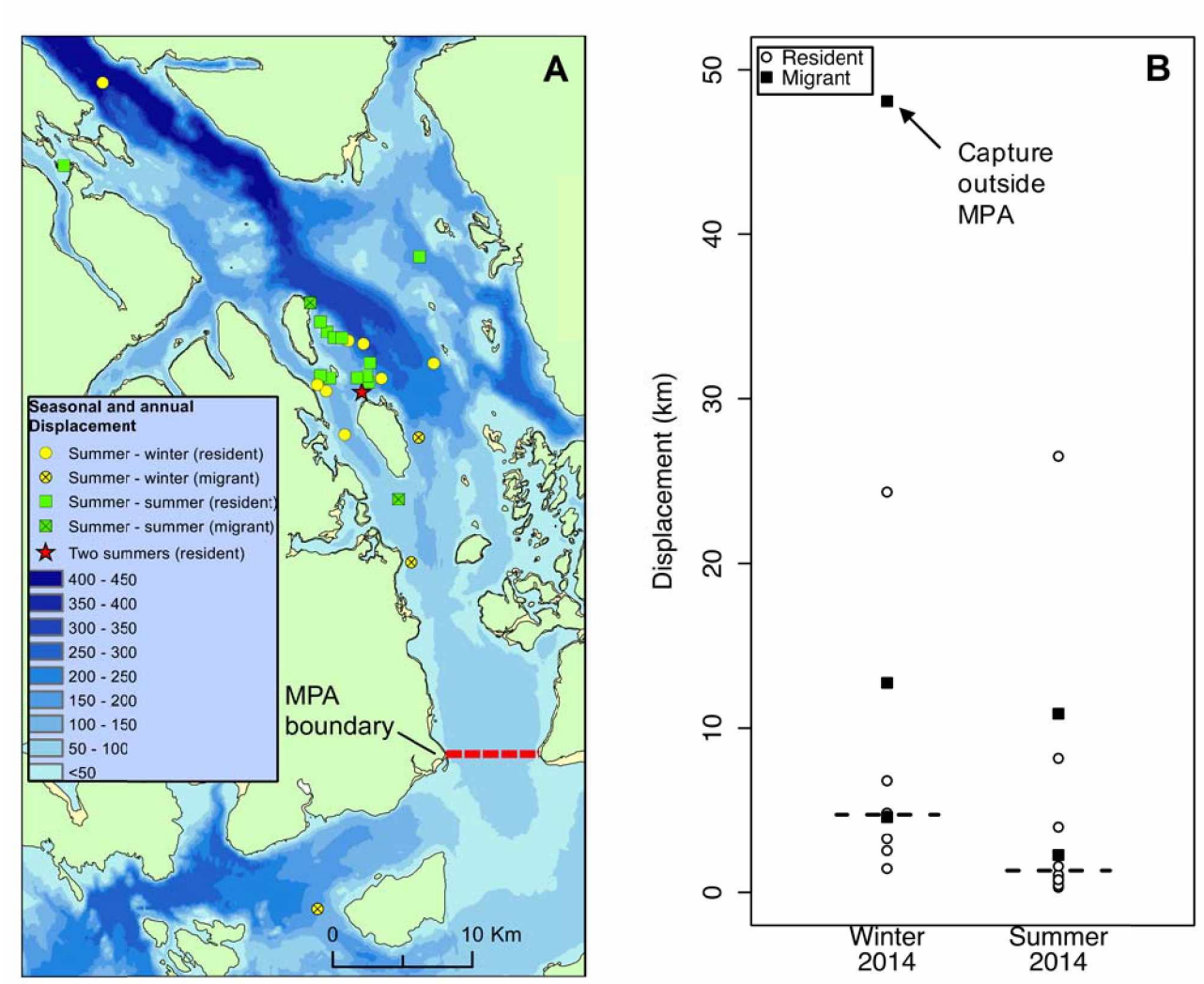


Figure 2.5. Seasonal and annual displacement of tagged halibut from release locations.
A) Locations of tagged fish during winter (February 1 pop-up date and fisheries recaptures; yellow circles) and the following summer (July 1 pop-up date, acoustic telemetry, and fisheries recaptures; green squares). Migrant fish are shown by hatched symbols. B) Seasonal displacement distance from release locations for summer-to-winter and summer-to-summer time periods. Resident fish are indicated by white circles and migratory fish by black squares. Dashed lines indicate the median displacement for each season.

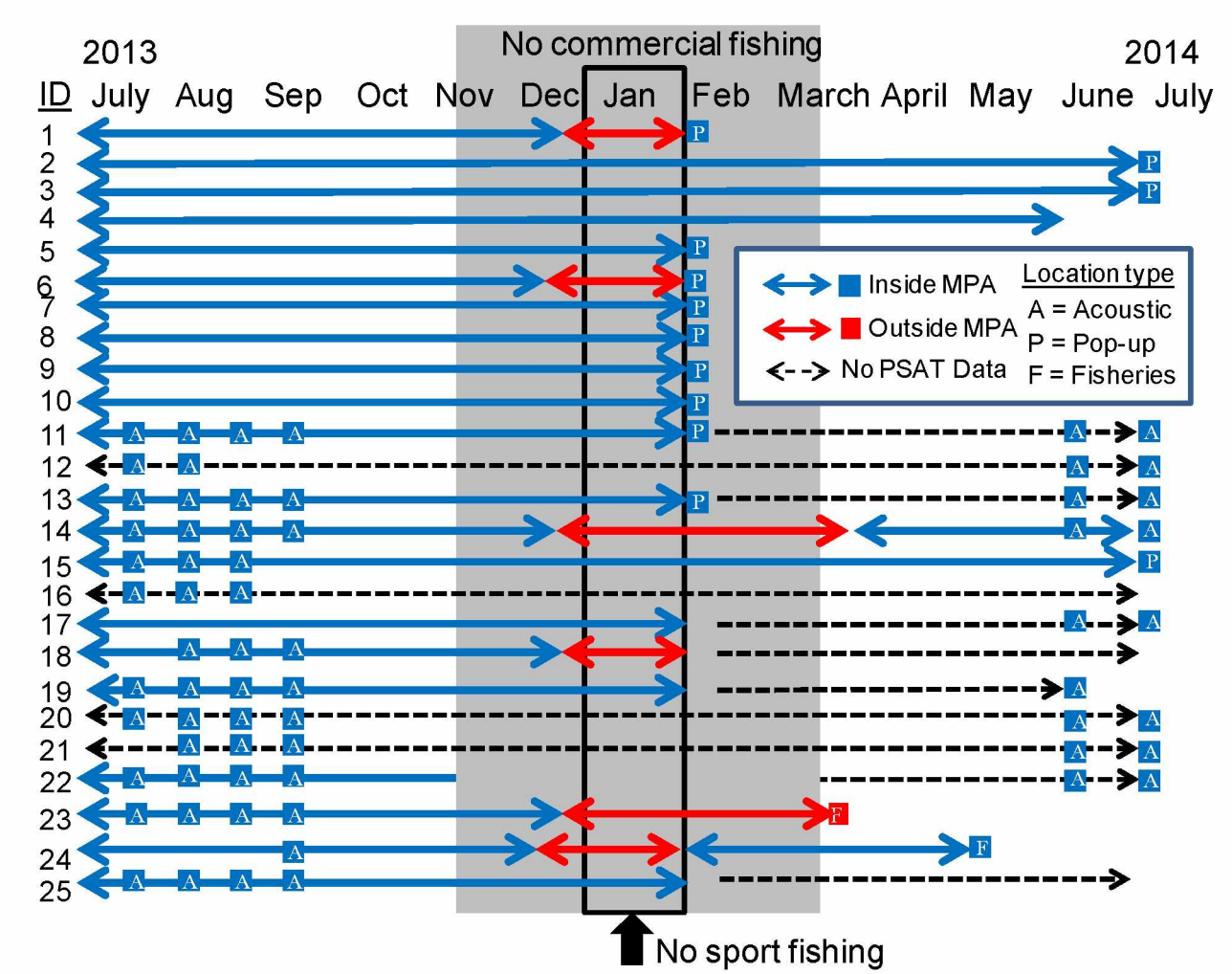


Figure 2.6. Summary of residence and migration of halibut based on PSAT data and acoustic tracking, PSAT pop-up, and fisheries recapture locations. Boxes denote locations known to be inside (blue) or outside (red) of the MPA. Blue lines denote periods of time when tagged fish were inferred to be inside the MPA, and red lines denote time inferred to be outside based on residence criteria. A dashed line indicates that PSAT data were not available. The larger gray box indicates the commercial fishery closure duration, and the smaller black rectangle indicates sport fishing closure. Note that only Fish #s 11 – 25 received acoustic tags.

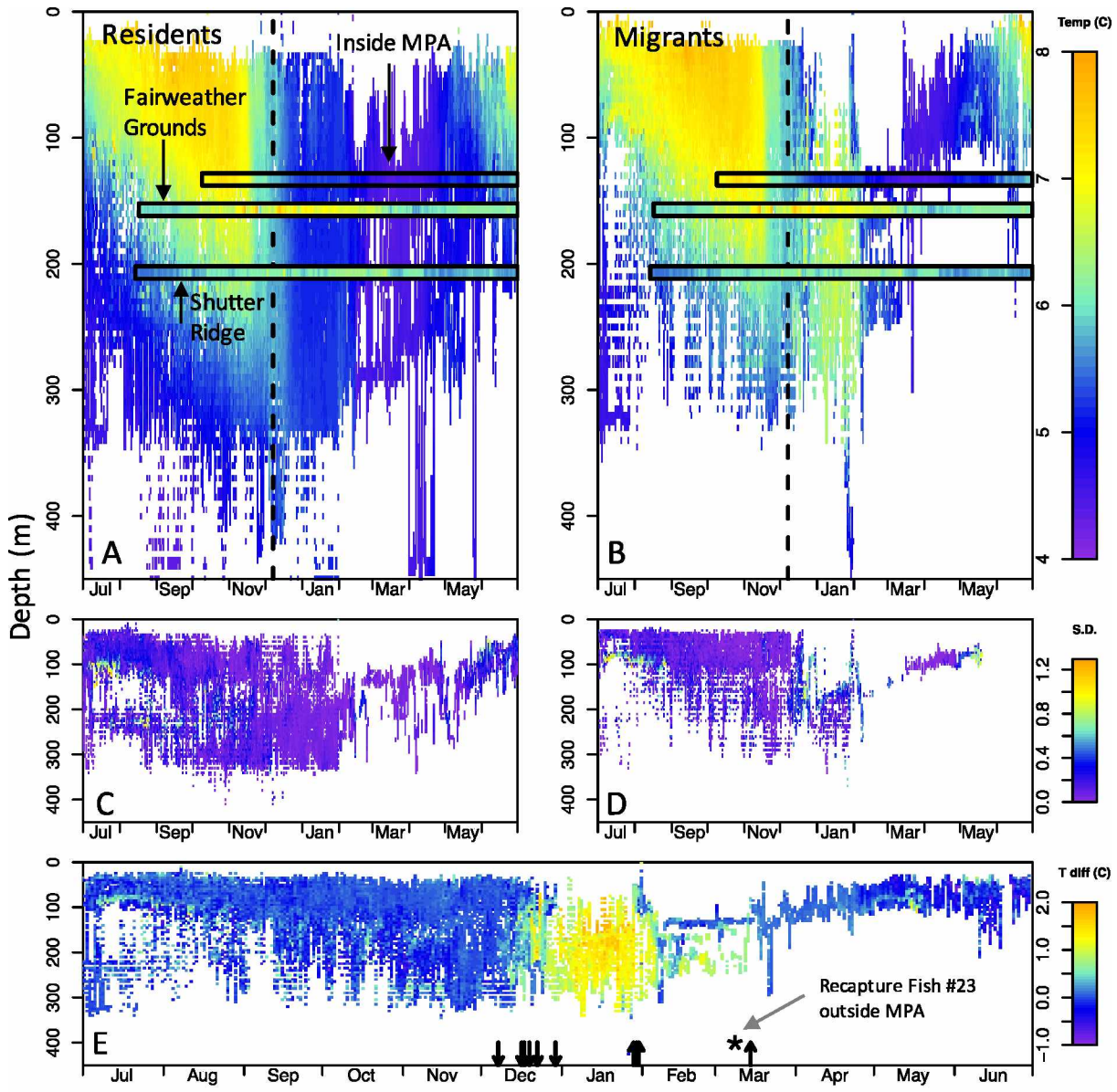


Figure 2.7. Composite TDPs in 5-m depth bins for A) resident fish (n = 12) and B) migratory fish (n = 6). Horizontal bars represent temperature-depth values from stationary moorings in the core study area (inside the MPA), the Fairweather Grounds (Gulf of Alaska), and Shutter Ridge (Gulf of Alaska). The first estimated date of departure for migrations is denoted by a dashed line on December 8. Standard deviation in temperature for each depth bin (for > 1 observation per depth bin) are shown in C (residents) and D (migrants). (E) Temperature differences between the migrant and resident. Positive values indicate that migratory fish experienced higher temperature than resident fish. Dates of inferred sill crossings on outward migrations are indicated by downward arrows and return migrations by upward arrows (exits and returns that occurred on the same day were jittered for visibility). Date of recapture for Fish #23, recaptured outside the MPA, is shown by an asterisk.

2.9 Appendices

2.9.1 Supplement S2-1: Acoustic tracking methodology and detection distance information

In this appendix, we provide additional details on fine-scale and high-speed tracking procedures and present information on detection distances typical for each method. Six tracking trips were conducted as part of this research (four in the summer of 2013, two in the summer of 2014). We present results from Tracking Trip 2 (August 3 and 4, 2013) as an example of the type of tracking data that were collected and typical detection distances that were achieved.

2.9.1.1 Tracking procedures

2.9.1.1.1 Test tag mooring

During each tracking trip, we deployed a temporary test tag mooring. The mooring was composed of a longline anchor, leaded line, and a surface buoy. We attached two transmitters to the line in a horizontal orientation (the long axis of the tag was perpendicular to the line). We attempted to place the tags at depths of 100 m and 50 m, however these depths varied from trip to trip. Within each trip, the location of the tags relative to the anchor varied slightly due to tidal and wave action on the surface buoy (e.g., the mooring line was pulled north of the anchor on an incoming tide and south on an outgoing tide). The transmitters for the test tag mooring for tracking trip 2 were deployed at depths of 56 m and 84 m. The water depth at the mooring location was 100 m.

2.9.1.1.2 Fine-scale tracking

During fine-scale tracking, estimated locations of tagged fish were obtained by navigating the vessel in the vicinity of the tagged fish using real-time information on the angle between the vessel and the tagged animal, signal strength, and the depth of the tagged fish. Fine-scale tracking was conducted using two Lotek LHP1 omni-directional hydrophones suspended from the bow on each side of the tracking vessel (Figure S1-1 A) and a 2-port Lotek Map 600 RT receiver. When a tag transmission was received by both hydrophones, the receiver calculated the direction of the transmitting tag relative to the orientation of the hydrophones in 10-degree increments. In addition to angular information, a measure of signal strength was also recorded by the receiver as a value between 0 (low) and 100 (maximum). In

general, signal strength decreased as distance between the tag and the hydrophones increased. Therefore, very high values indicated that the vessel was close to the tagged fish. However, because transmissions close to the hydrophone can also have a low signal strength, low values did not necessarily indicate that the tagged fish was far away. Additional clues to the location of tagged animals were provided by the acoustic transmitters that transmitted the depth of the fish in 7-m bins. The depth of the fish was matched to bathymetric maps of the study area to narrow down potential fish locations. Using a laptop equipped with ArcGIS, we monitored the vessel trajectory, angles to the tagged fish, trends in signal strength, and matched the depth of the fish to bathymetry grids to estimate locations of tagged fish along with estimated error. The primary information used for estimating locations was triangulation of angles between the vessel and tagged animals from different vessel locations, and we sought to maneuver the vessel in a circle around the tagged fish to confirm its location when possible. For this tracking method, the primary function of the test tag mooring was to ensure that the real-time directional information provided by the receiver was correct.

2.9.1.3 High-speed tracking

High-speed tracking was performed to obtain coarse-scale information on the location of tagged animals over a larger area than the fine-scale tracking allowed. In this tracking configuration, a single hydrophone was attached to a v-fin and towed behind the vessel (Figure S1-1 B) at speeds of 5 – 8 knots. On each tracking trip, regular transects were conducted within the core area on a 400-m grid assuming a 200-m detection distance minimum. For this tracking method, the primary function of the test tag mooring was to choose a vessel speed that would provide the necessary detection distance prior to initiating transects. Information on maximum signal strength vs. distance was also used to estimate distance of the tagged animal from the transect line.

2.9.2 Results from Tracking Trip 2

2.9.2.1 Fine-scale tracking

For the fine-scale tracking configuration, the 56-m test tag was detected at a maximum distance of 656 m and the 84-m tag at a maximum distance of 630 m. Most signal strength values > 70 occurred at

distances < 80 m from the 56-m tag and < 30 m from the 84-m tag (Figure S2.1-2). In general, at distances less than 200 m both tags were detected frequently. Signal strength was stronger for the 56-m tag (Figure S2.1-3 A) compared to the 84-m tag (Figure S2.1-3 B). It is estimated that tagged fish in the core study area were detected over similar spatial extents compared to the test tags (Figure S2.1-4).

An example of angular and signal strength information provided by fine-scale tracking is provided by detection data for the 54-m test tag (Figure S2.1-5). The vessel maneuvered around the test tag mooring for approximately 35 minutes. Because the tracking occurred during an outgoing tide, it is likely that the tag was slightly south of the anchor location.

2.9.2.2 High-speed tracking

The maximum distance at which the 56-m test tag was detected with the v-fin configuration was 460 m compared to a maximum distance of 275 m for the 84-m tag. As was the case for fine-scale tracking, signal strength was greater for detections of the 56-m tag compared to the 84-m tag (Figure S2.1-6). Note that the goal of the vessel maneuvers around the test tag was to ensure detections at a distance of 200 m rather than to perform explicit range tests to determine detection limits. However, the diminishing signal strength around 300 m for both test tags suggests that 300 m may be close to a maximum detection limit. Both test tags were detected at distances > 250 m at speeds of 5 – 7 knots (Figure S2.1-7). A 250-m buffer of the vessel path during the transects of the core study area suggests that tracking with the v-fin allows comprehensive coverage of the core study area for depths less than approximately 100 m (Figure S2.1-8). The core study area is located on a mound with most depths less than 100 m, so although detection distance is likely to be smaller in deeper waters on the edges, the majority of the core study was adequately covered.

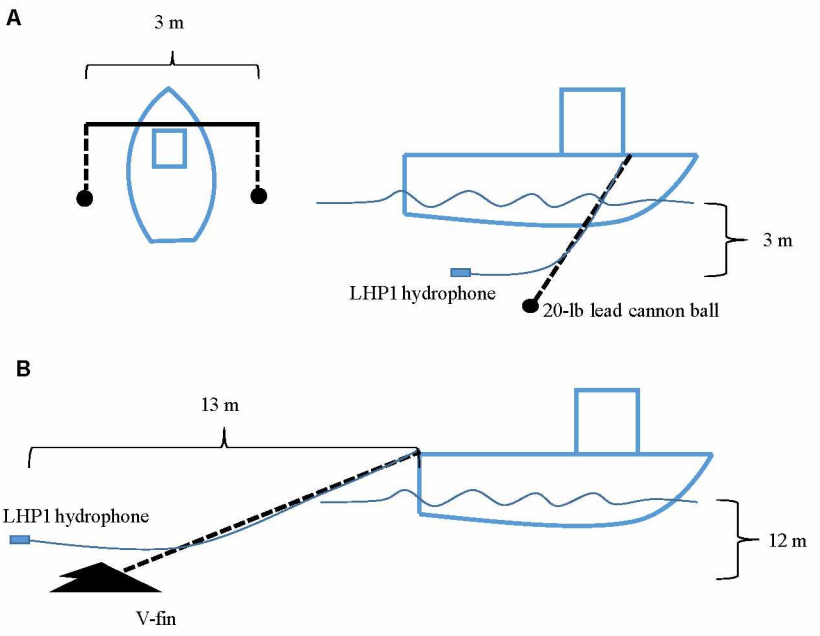


Figure S2.1-1. Hydrophone configurations for A) fine-scale tracking and B) high-speed tracking with the v-fin.

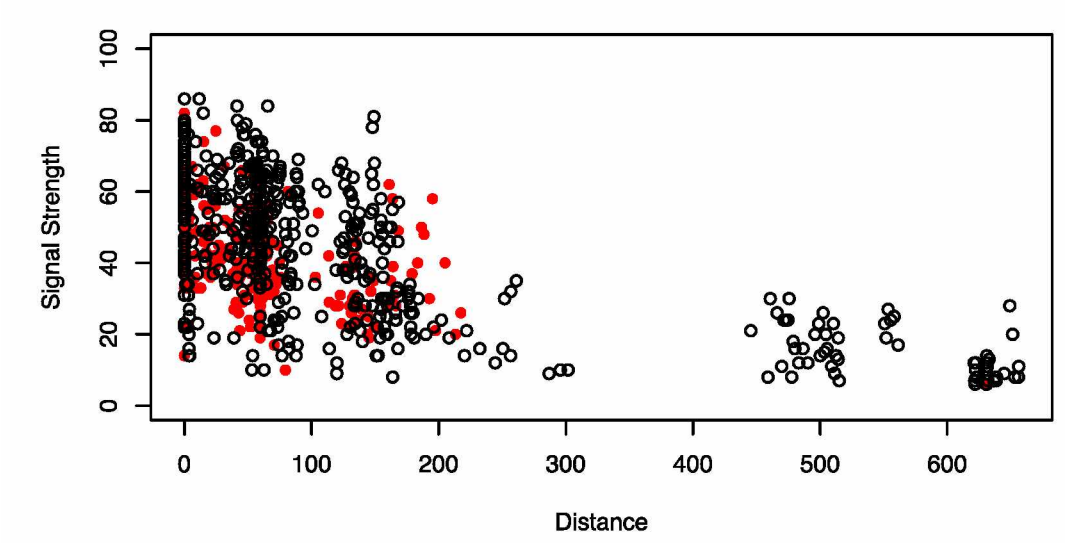


Figure S2.1-2. Signal strength vs. distance from 56 m (open back circles) and 84 m (solid red circles) test tags during fine-scale tracking on Tracking Trip 2.

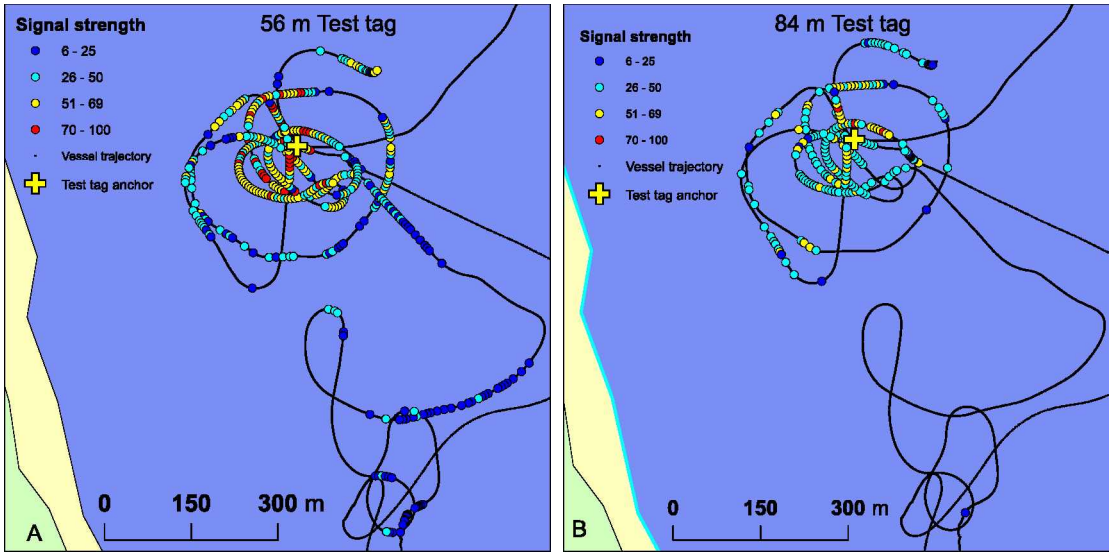


Figure S2.1-3. Fine-scale tracking detection of the 56-m test tag (A) and the 84-m test tag (B). Detections are color coded by signal strength.

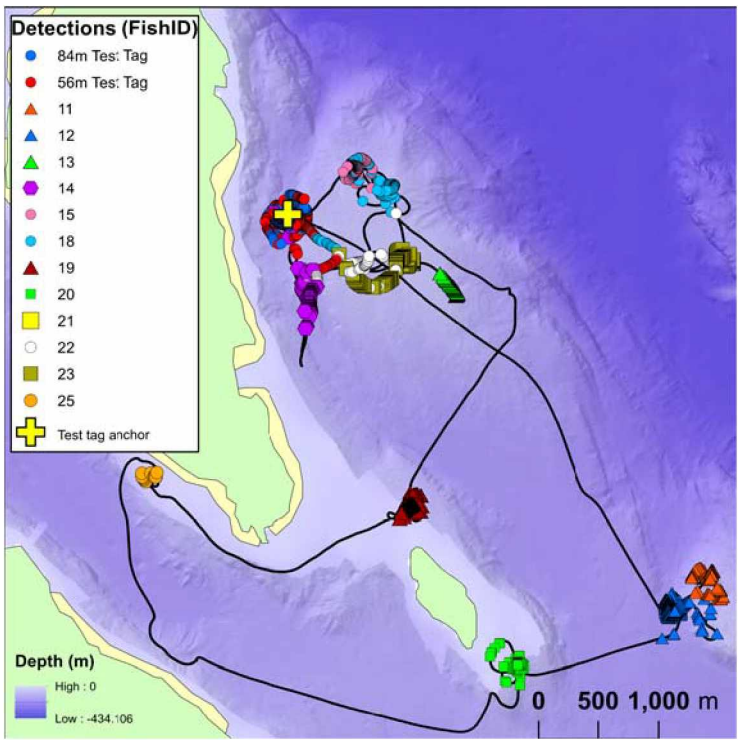


Figure S2.1-4. Detections of test tags and tagged fish during fine-scale tracking in the core study area for Tracking Trip 2.

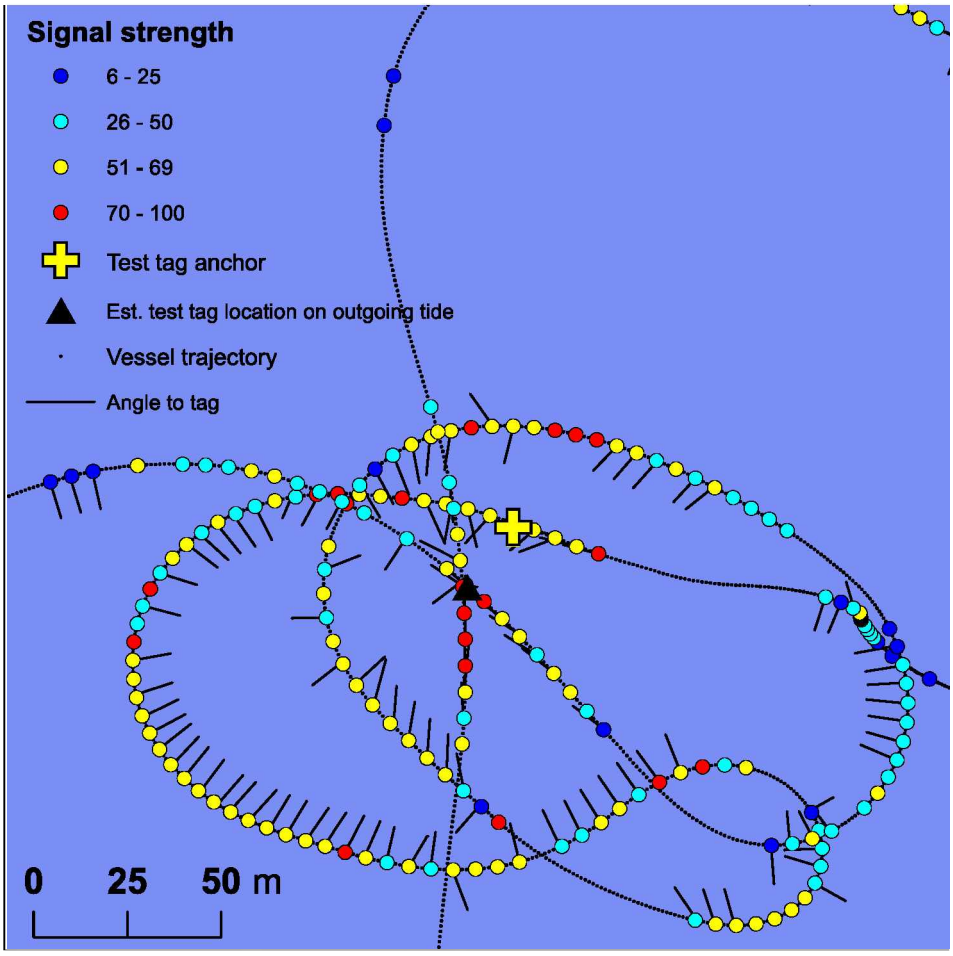


Figure S2.1-5. Fine-scale tracking data for the 56-m test tag. Detections are color coded by signal strength. The estimated angle to tagged animal (available when a signal was received at both hydrophones) is indicated by thin black lines. Detections occurred over a time period of approximately 35 minutes. The yellow cross indicates the location of the test tag anchor, and the black triangle indicates the possible position of the test tag on the line during the outgoing tide.

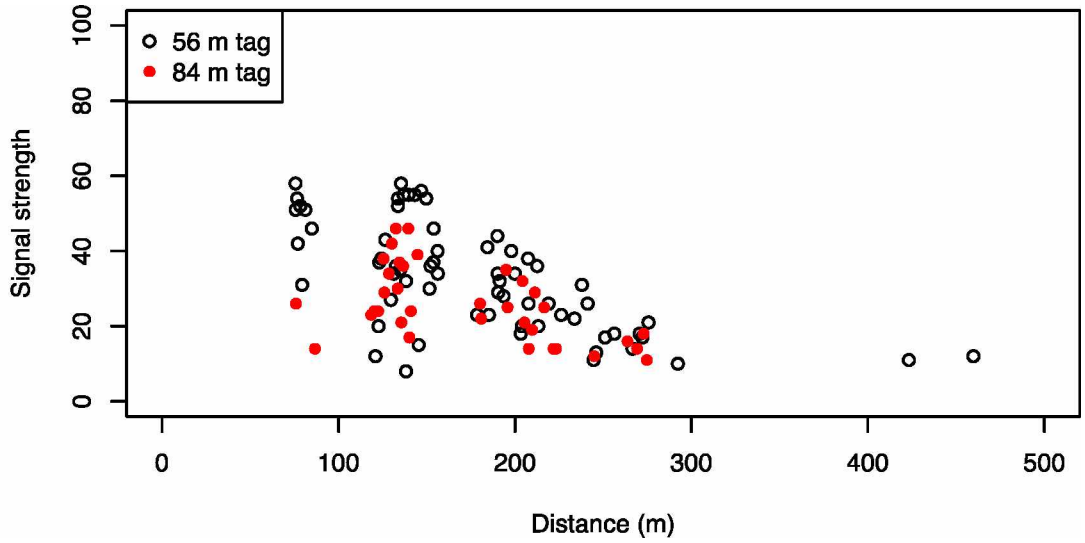


Figure S2.1-6. Signal strength vs. distance from 56-m (open back circles) and 84-m (solid red circles) test tags during high-speed tracking with the v-fin on Tracking Trip 2.

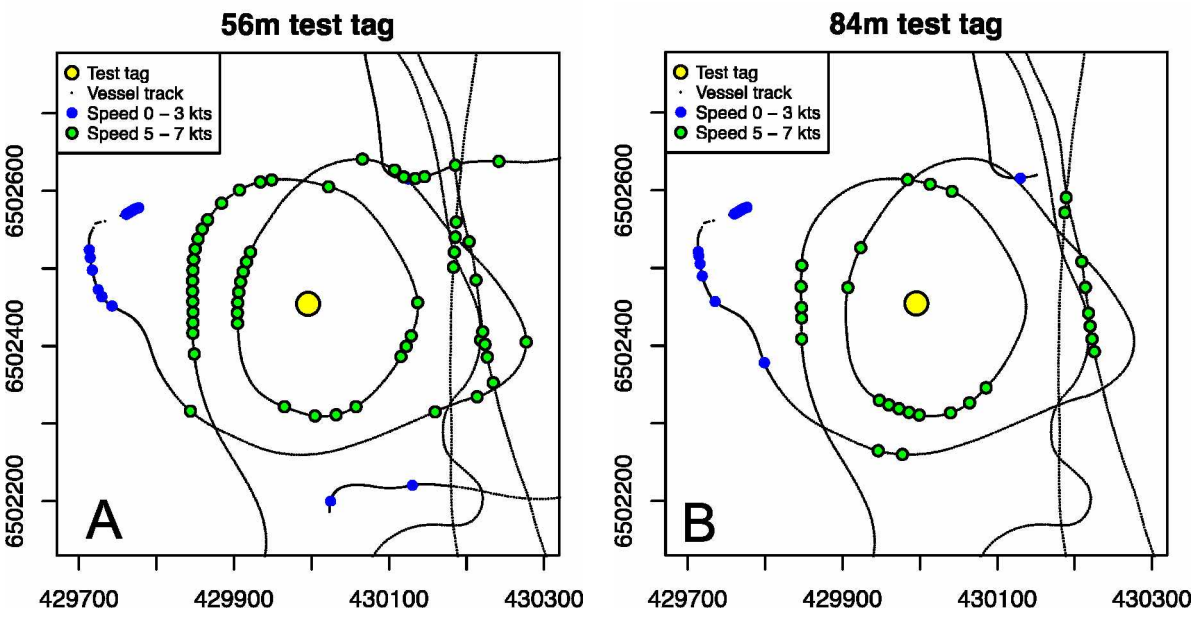


Figure S2.1-7. Detections of (A) the 56-m test tag and (B) the 84-m test tag during high-speed tracking with the v-fin configuration. Detections obtained while vessel speeds were > 5 knots are shown in green, detections at speeds < 3 knots are shown in blue.

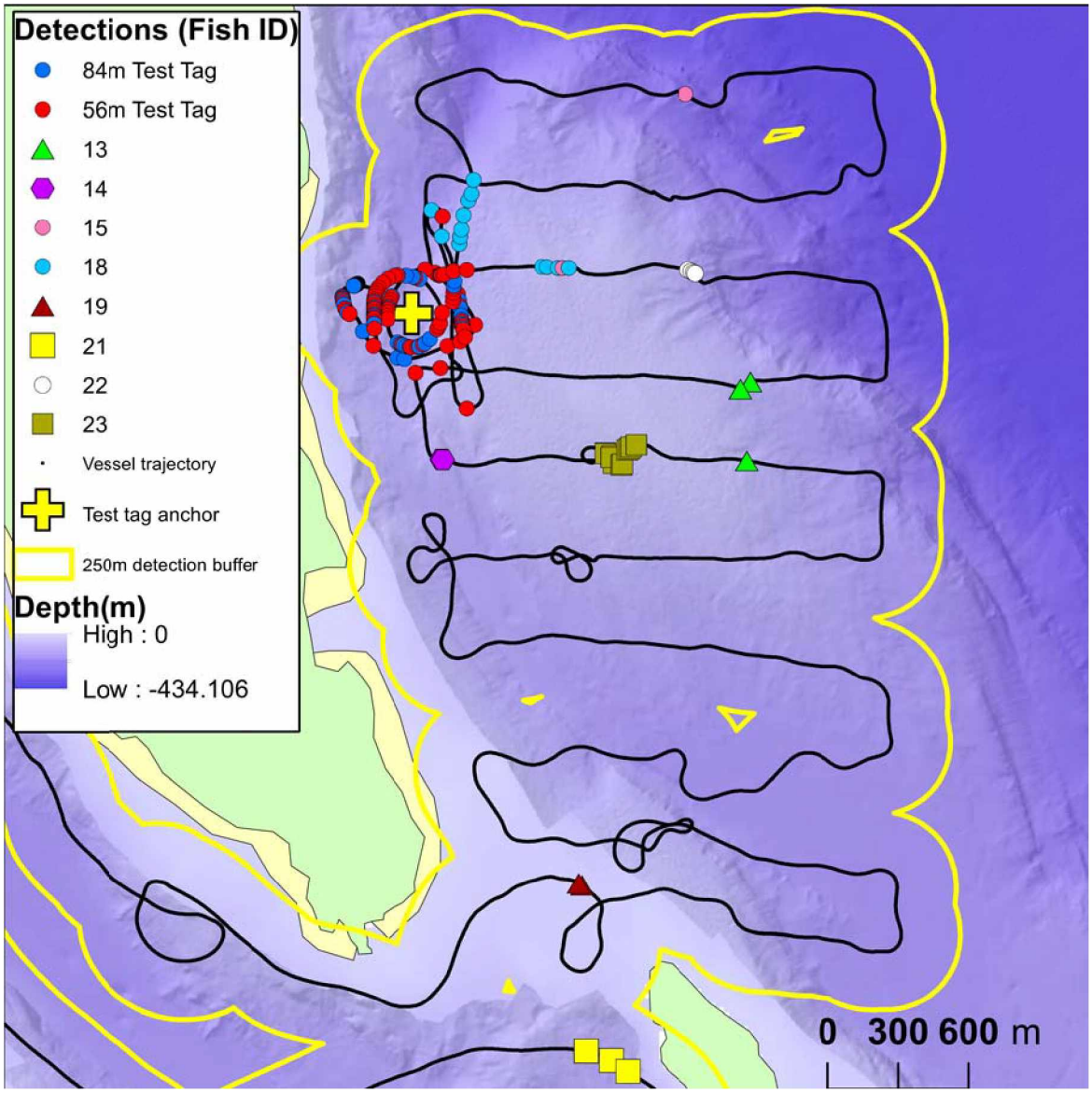


Figure S2.1-8. Vessel trajectory during high-speed tracking along a 400-m grid of the core study area during Tracking Trip 2. A 250-m buffer around the vessel path (yellow line) representing detection distance indicates coverage of the core area. Detections of individual tagged animals and test tags are shown by solid symbols.

2.9.2 Supplement S2-2: Procedure for inferring MPA residence and migration from PSAT data

We used a combination of known locations (from acoustic telemetry, PSAT pop-ups, and fisheries recaptures) and PSAT data to infer whether tagged fish were inside or outside of the MPA for each day of their archival tag records. Criteria used to classify fish as "residents" or "migrants" consisted of 1) a maximum daily depth threshold that reflects dates on which fish may have crossed the shallow sill at the entrance to the MPA, 2) a maximum temperature threshold that precludes winter residency within the MPA, 3) matching PSAT temperature-depth profiles to data provided by the stationary mooring and multiple oceanographic surveys within the MPA and two stationary moorings in the Gulf of Alaska, and 4) matching tidal amplitude and phase from PSAT depth records to the stationary mooring in the core study area. The number of criteria available for individual data sets varied depending on PSAT tag type (either Wildlife Computers or Desert Star) and data type (either summarized data from Argos transmissions or high-resolution data from physically-recovered tags).

The first criterion (applied to all PSAT records) was whether or not the maximum daily depth recorded by a tagged fish was ever shallower than the depth of the glacial sill at the entrance of the MPA. Water depths range to 450 m within the MPA, but the glacial sill at the entrance to the fjord has created a ring of shallow (40 m) water that approximates the management boundary (Figure 1B). Departure of the tagged fish from the MPA results in recorded depth values shallower than 40 m. We developed a maximum daily depth threshold criterion to determine possible dates of tagged fish exit from (and re-entry to, if applicable) the MPA. The maximum daily depth criterion was conservatively set at 70 m to allow for uncertainties in depth accuracy of the tag (± 20 m offset for DS tags), tidal stage (7 m range), and accuracy of the bathymetric grid. Because resident fish regularly occupied depths less than 70 m, satisfying this criterion alone was not deemed to be sufficient evidence of movement across the sill; rather, this was used in combination with other criteria (described below). As a corollary, tagged fish with maximum daily depth records that were always in excess of 70 m for the entire PSAT record were automatically classified as residents.

A second criterion for MPA residency consisted of a maximum winter temperature threshold (applicable to all PSAT data sets). This criterion was used to identify fish that could not have resided within the MPA during January and February based on PSAT maximum temperatures that were warmer than a threshold temperature in the MPA during this time period. Two decades of oceanographic sampling have revealed seasonal patterns of water column temperature throughout the MPA (Johnson and Sharman, 2014). During the summer months the water column is stratified and temperature-at-depth increases gradually from April – November. The water column becomes well-mixed during December and then the entire water column cools from January – March, when very little spatial variation in temperature is typically observed throughout the bay (Danielson, 2012). Based on an analysis of the NPS historical oceanographic data set, the maximum water temperature expected anywhere within the MPA during January or February 2014 is 5.5°C (Supplement S2-3). Therefore, fish that recorded temperatures greater than 5.5°C during January or February were inferred to be located outside of the MPA during that time period.

A third criterion (applicable to all WC PSATs and all detailed data sets) assumes that fish with temperature-depth profiles that match temperature and depth data from the stationary mooring and oceanographic survey station in the core study area are residents of the MPA. Temperature-depth profiles, or vertical profiles of water column temperature at depth (referred to subsequently as "TDPs"), are characteristic of specific water bodies and thus have previously been used for geolocation of fish tagged with PSATs (Skomal et al., 2009). TDPs are known to vary spatially, particularly in fjord-type estuaries, where different water bodies have been observed to possess distinct TDPs (Moore et al., 2008). Glacier Bay has unique oceanographic properties that result from fjord topography (shallow sill at the mouth and deep basins), freshwater runoff, sedimentation and cold temperatures associated with glacial runoff (Etherington et al., 2007). Therefore, the oceanography of Glacier Bay differs from adjacent waters of Icy Strait, Cross Sound, and the Gulf of Alaska (Danielson, 2012).

TDP information was available from detailed DS and WC data sets and summarized WC data sets. For detailed tags, TDPs were created for each day of the archival tag record by calculating the maximum

temperature within 5 m depth bins that ranged between 0 and 450 m. Summarized WC data provided TDP information as the maximum and minimum temperatures within each of 8 depth bins that equally spanned the range of depths recorded in each 12-hour time bin. To allow comparisons with daily TDPs from detailed data sets, depth bins from summarized data were synchronized to the same 5 m bins as the detailed tags and the maximum temperature in each 5-m bin from both 12-hour bins from each day combined was selected for the summarized TDP.

We matched TDP data from tagged fish to stationary mooring and CTD data from Oceanographic Survey Station 4 by comparing fish TDP data to oceanographic and stationary mooring data obtained inside the core study area. We then overlaid the oceanography TDP data from CTD casts at Station 4 (8 surveys were conducted over the course of the study) and data from the stationary mooring in the core study area (average daily data at a mean depth of 132 m from October 2013 – July 2014). CTD and mooring data were formatted to match the tagged fish TDP data by calculating the maximum temperature in 5 m depth bins that ranged between 0 and 450 m. This resulted in vertical columns for CTD data (one for each day surveyed) and horizontal rows for mooring data (continuous over time but in one depth bin). We visually determined whether profiles for the fish and the study area were qualitatively similar. Days when PSAT TDPs matched core study area TDPs were classified as "residential".

To determine the spatial extent over which the PSAT TDPs that matched TDPs from the core study area also matched oceanographic data collected at other stations within and adjacent to the MPA, we compared halibut TDP data to oceanographic data from seven survey stations located throughout the MPA and one station located outside the MPA (Figure 2.1B). Oceanographic surveys were conducted at these eight stations during July, August, September, October, March, April, early May, and late May. For each oceanographic survey, we plotted the temperature from CTD casts (1-m intervals) at eight oceanographic stations. We then overlaid fish TDP data plotted as a range of temperatures recorded in 1-m depth bins during the 24-hour period of the survey. We also added the range of temperatures observed for the stationary mooring in the core study area during the 24-hour period in which Oceanographic Station 4 was sampled. In this way, it was possible to visualize whether TDPs from tags attached to fish

were more similar to stations closest to glaciers (Stations 12, 16, and 20), in the central bay (Stations 4, 7, and 13), at the mouth of the MPA (Station 1), or adjacent to the MPA in Icy Strait (Station 24) at different times of the year.

Comparisons of fish TDP data to oceanographic data collected throughout and adjacent to the MPA during eight oceanographic surveys over the course of the study provided additional confirmation of residence and migration classification for tagged fish. For example, comparison of TDPs for Fish #23, #15, and #14 to multiple oceanographic stations during different oceanographic surveys indicated that during the July and October 2013 surveys, fish TDPs more closely matched the survey stations in the central portion of the MPA rather than the head of the bay, at the entrance, or outside the MPA (Figure S2.2-1, A and B). However, during the March 2014 survey, the TDP of the resident Fish (#15) matched CTD cast and stationary mooring data in the central portion of the MPA while both Fish #23 and #14 were more consistent with the TDP for Icy Strait station (located outside of the MPA, Figure S2.2-1, C). Fish #23 was recaptured in the fishery 12 km from the Icy Strait oceanographic survey station (#24) on the same day of the survey. During May, after fish #14 was inferred to have returned to the MPA, both Fish #14 and #15 had TDPs that matched stations in the central portion of the bay again (Figure S2.2-1, D). These results support a classification of "MPA resident" when PSAT TDPs matched TDPs from the stationary mooring and oceanographic data in the core study area.

A fourth criterion (applicable to all detailed data sets) assumes that when tidal amplitude and phase from PSAT depth records matched the tidal amplitude and phase recorded by the stationary mooring in the MPA, the tagged fish was likely to be located in the vicinity of the stationary mooring. The rise and fall of the tides can be detected in detailed depth records of stationary tagged fish. Because the time of high tide (phase) and tidal amplitude vary over space and can be modeled, comparing the tidal amplitude and phase measured from archival tags to predicted values is a common method of geolocation for archival-tagged demersal fish (Hunter et al., 2003). Differences between tidal amplitude and phase are large between Glacier Bay and adjacent locations (Shi et al., 2014).

We used detailed depth records from physically-recovered tags to perform a crude version of geolocation by comparing tidal amplitude represented as change in depth from the PSAT while the tagged fish was stationary to change in depth recorded at the stationary mooring inside the core study area. We manually offset depth records from the stationary mooring tags so that they matched the depth of the fish. If PSAT data and stationary mooring depth data overlapped, residency in the core study area was inferred. However, if they were offset in terms of amplitude or phase, the tagged fish was assumed to be outside of the core study area.

The procedure for inferring MPA residence or migration for each day of the archival tag record began with knowledge of dates when tagged fish were known with certainty (from acoustic tracking, PSAT pop-ups, or fisheries recaptures) to be either inside or outside of the MPA. The maximum daily depth threshold criterion was applied to determine potential dates of exit from or re-entry into the MPA. Then the criteria derived from comparing tag data to known depth and temperature inside and outside the MPA were employed to interpolate the status of residence/migration for days when known locations were not available. For example, a fish captured outside of the MPA that did not have a maximum daily depth less than 70 m for the previous month must have been outside the MPA for at least one month. For all tagged fish, if winter temperatures were warmer than the maximum temperature criterion, days when PSAT records differed from the core study area were classified as "migratory". For fish with TDP data, days during which PSAT TDPs matched stationary mooring and oceanographic survey data from the core study area were classified as "residential". For fish with detailed data sets, tidal amplitude/phase from PSAT data were compared to tidal amplitude/phase in the core study area for time periods between potential sill crossings and known locations.

We provide three examples of the procedure for inferring residence and migration for fish with PSAT data. The first example features migratory Fish #23, which was the only tagged fish to have an end location outside of the MPA. During the summer and fall, this fish exhibited a very small home range in the release location and was detected on every tracking trip. During this time, temperature-depth profiles (TDPs) from the PSAT data closely matched CTD casts from oceanographic surveys in the core study

area until late December, when they began to diverge from the stationary mooring data (Figure 2.2A). On December 30, the fish recorded a maximum daily depth that was less than the 70 m depth, meeting the criteria for crossing the shallow sill at the entrance to Glacier Bay. Because the fish was never shallow enough to cross the sill again before it was recaptured outside of the MPA in March, 2014, the fish must have exited the MPA no later than December 30. While the fish was known to be outside of the MPA, it experienced January temperatures that were greater than the criterion for winter residence within the MPA (5.5°C). Tidal amplitude/phase in the PSAT depth record matched the change in depth measured by the stationary mooring in the study area before December 30, but not after. Therefore, we infer that this fish likely left the MPA on December 30.

The second example features Fish #15, which was inferred to have resided inside the MPA throughout the winter. This fish exhibited home range behavior in the release location during the summer of 2013. Although the fish maximum daily depth was shallower than the depth threshold in October and again in June, TDPs from the PSAT data matched both CTD casts from the central bay and stationary mooring data throughout the winter (Figure 2.2B) and tidal amplitude/phase in the PSAT record matched conditions in the core study area year round. The pop-up location in July 2014 was 8.2 km from the release location. Based on these lines of evidence we infer that the fish likely resided within the MPA for the entire year.

The third example features Fish #14, which was found to undertake a seasonal migration outside of the MPA in December but returned to the MPA in March. TDPs from PSAT data matched CTD casts from the central bay and stationary mooring data until December 17, when TDPs became warmer than the stationary mooring data in conjunction with the fish attaining sufficiently shallow depth to allow it to cross the sill at the entrance to the MPA (Figure 2.2C). During January, temperatures at depths > 100 m were almost 2°C warmer than the criterion for winter residence within the MPA (5.5°C). Five putative spawning rises were observed between January 17 – 29. On March 15, the maximum daily depth of the fish was again shallow enough to allow the fish to cross the sill at the MPA entrance, after which the TDPs matched the temperature depth profiles of both CTD casts and the stationary mooring data in the

core study area. The fish was located with acoustic tracking within 1 km of the release location during June 2014. Tidal amplitude/phase matched stationary mooring data before December 17 and after March 15, but during January the amplitude and phase differed from the MPA. Thus, we conclude that this fish left the MPA during December to spawn at a location outside of the MPA, but following the spawning migration returned to the same small home range that it had occupied in 2013.

2.9.2.1 References

Danielson, S. L. 2012. Glacier Bay oceanographic monitoring program analysis of observations, 1993–2009. Natural Resource Technical Report NPS/SEAN/NRTR—2012/527. National Park Service, Fort Collins, Colorado. 172 pp.

Etherington, L. L., Hooge, P. N., Hooge, E. R., and Hill, D. F. 2007. Oceanography of Glacier Bay, Alaska: implications for biological patterns in a glacial fjord estuary. Estuaries and Coasts, 30: 927-944.

Hunter, E., Aldridge, J. N., Metcalfe, J. D., and Arnold, G. P. 2003. Geolocation of free-ranging fish on the European continental shelf as determined from environmental variables: I. Tidal location method. Marine Biology, 142: 601-609.

Johnson, W. F., and Sharman, L. C. 2014. Glacier Bay National Park and Preserve oceanographic monitoring protocol: Version OC-2014.1. Natural Resource Report NPS/SEAN/NRR — 2014/851, National Park Service, Fort Collins, Colorado. 259 pp.

Moore, S. K., Mantua, N. J., Newton, J. A., Kawase, M., Warner, M. J., and Kellogg, J. R. 2008. A descriptive analysis of temporal and spatial patterns of variability in Puget Sound oceanographic properties. Estuarine Coastal and Shelf Science, 80: 545-554.

Shi, L., Wang, J., Myers, E., and Huang, L. 2014. Development and use of tide models in Alaska supporting VDatum and hydrographic surveying. Journal of Marine Science and Engineering, 2: 171-193.

Skomal, G. B., Zeeman, S. I., Chisholm, J. H., Summers, E. L., Walsh, H. J., McMahon, K. W., and Thorrold, S. R. 2009. Transequatorial Migrations by Basking Sharks in the Western Atlantic Ocean. Current Biology, 19: 1019-1022.

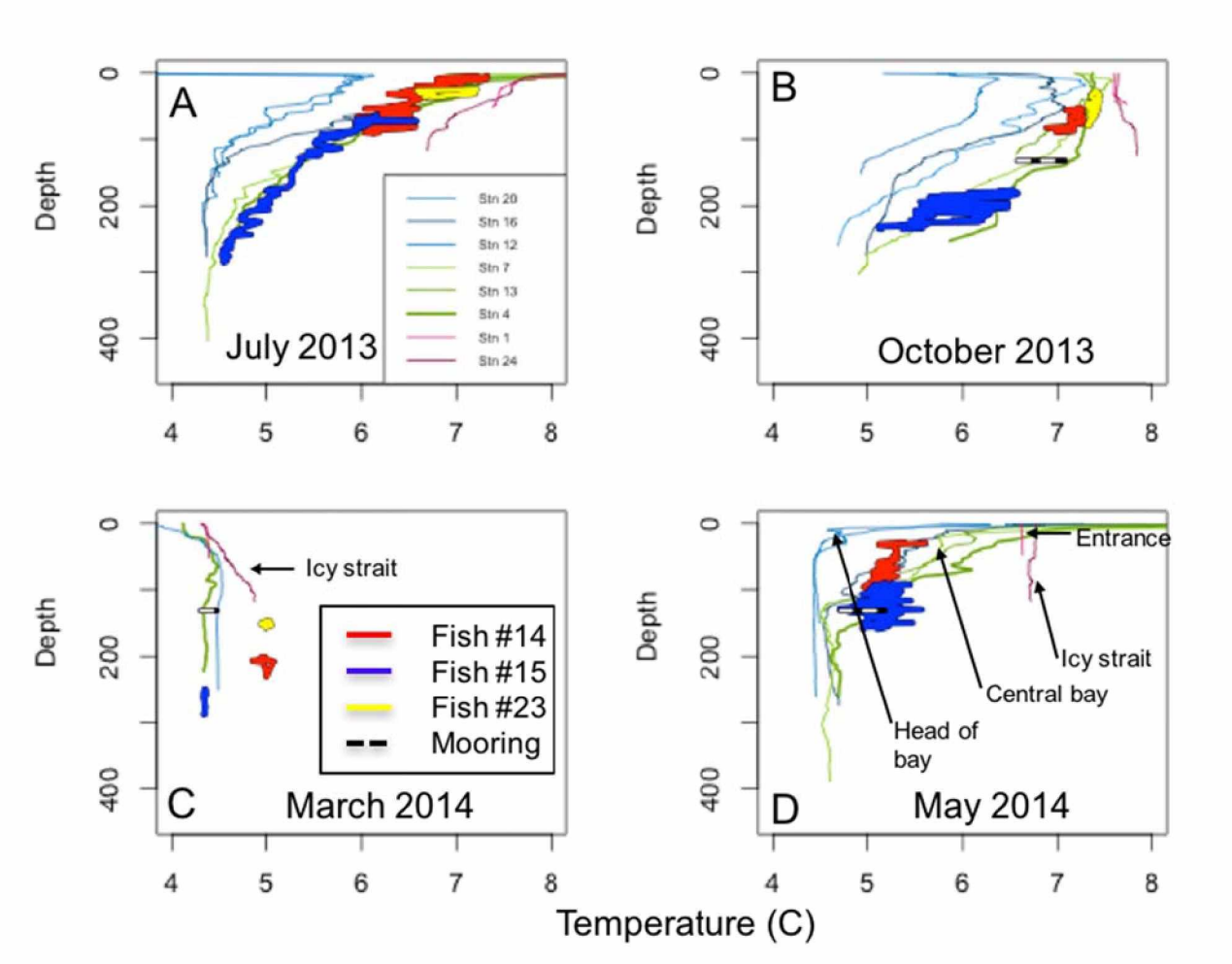


Figure S2.2-1. Example of daily comparison of PSAT data and CTD profiles at core stations during A) July, B) October, C) March, and D) May. Stations at the head of the bay (20, 16, and 12) are shown by thin lines in blue shades, stations in the center of the bay (4, 7, and 13) are shown in green shades, and stations at the mouth and Icy Strait (1 and 24) are shown in shades of pink. Temperature-depth profiles for tagged fish on the day of each CTD cast are shown for: Fish #23 (thick yellow line), which left the MPA in December based on the maximum daily depth criterion and was recaptured in Icy Strait several km from the CTD station on the same day that the CTD cast was taken; Fish #14 (thick red line), which left the MPA in December and returned in mid-March; and Fish #15 (thick blue line), which was remained in the MPA throughout the winter (TDPs for these three fish are also shown in Figure 2). The range of temperatures observed at the stationary mooring in the study area in October, March, and May is shown by a thick black and white dashed line.

2.9.3 Supplement S2-3: Developing a maximum winter temperature threshold for inferring MPA residence

We conducted a historical analysis of Glacier Bay oceanographic survey data to determine the maximum temperature that would be expected within the MPA during January and February of 2014. First, we examined the data set to determine whether winter trends in temperature-depth profiles by year were present. We plotted temperature vs. depth in January or February (depending on when the winter sampling trip was conducted each year) for all years and all stations (Figure S2.3-1). From this plot it is apparent that years 1994 – 2003 (solid lines) were warmer than years 2005 – 2013 (dashed lines). Also, at depths greater than 100 m, there was very little change in temperature with increasing depth.

To establish whether the winter of 2014 was cold relative to the winters for the years depicted in the graphic, we compared the temperatures observed at NPS Oceanographic Station 4 (in the study area) at a depth of 150 m to the monthly range of temperatures observed from the stationary mooring in the study area (Figure S2.3-2). During December, January, and February, the average values from the stationary mooring data were lower than the average temperatures from the CTD data, and the upper ranges of the stationary mooring data did not exceed the highest values from the CTD casts. Therefore, we conclude that January 2014 could be considered a cold year.

To establish a maximum temperature criterion for residency of halibut within the MPA during the winter of 2014, we calculated the mean and range of the maximum temperature observed at all stations during cold years (2005 – 2014) at depths greater than 100 m. The mean temperature was 4.9 °C, with a maximum observed temperature of 5.5 °C (Figure S2.3-3). Thus, although fish that resided in Glacier Bay would be expected to experience maximum temperatures closer to 5°C in January or February, we used the maximum temperature to serve as a conservative threshold for inferring MPA residency (e.g. fish that record temperatures higher than 5.5°C at depths > 100 m during January and February were not likely to be located anywhere within the MPA).

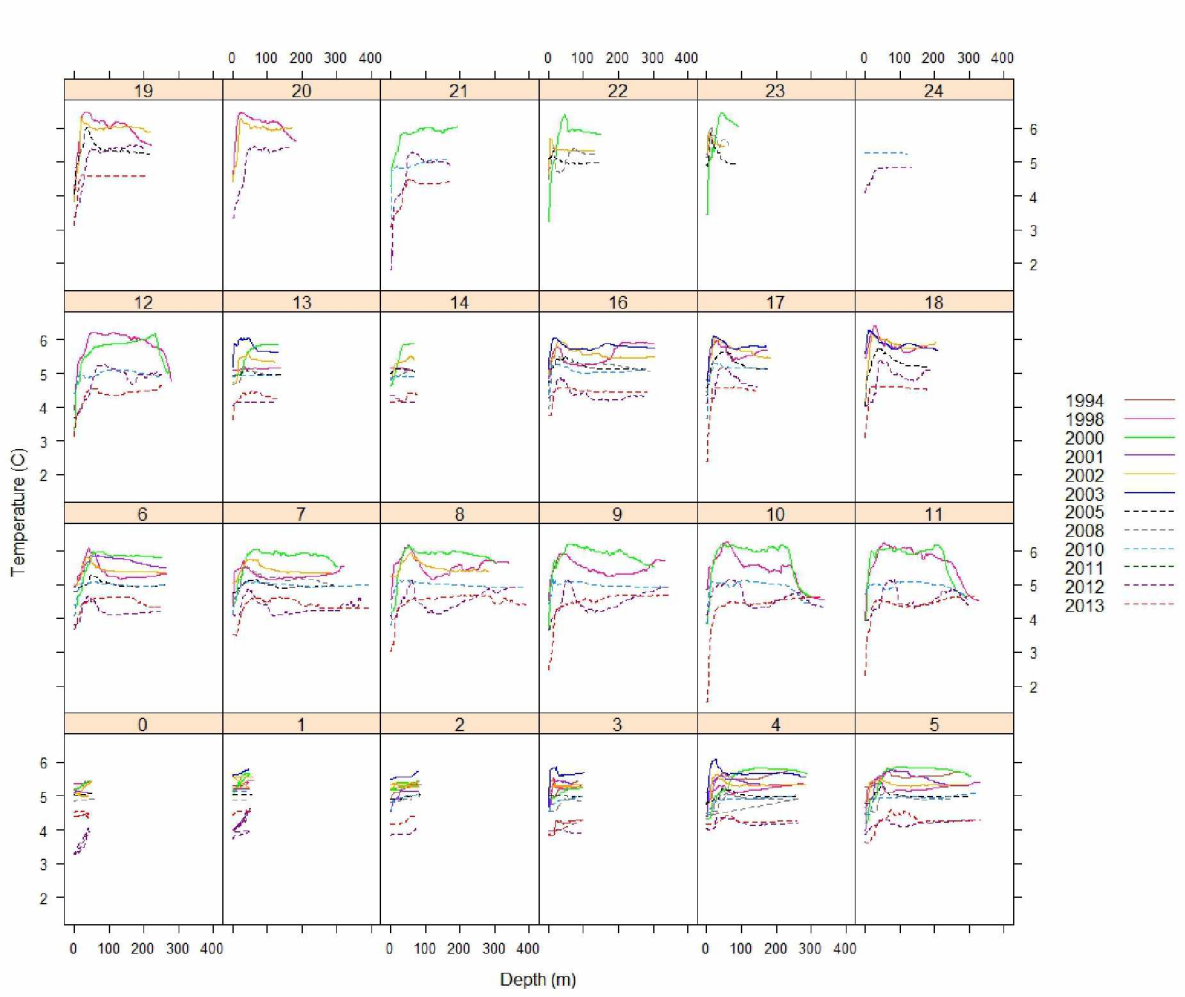


Figure S2.3-1. Temperature-depth profiles from NPS oceanographic surveys for all stations (0 to 24) in all years (1994 – 2013) during either January or February. Note that in some years, surveys were not conducted in either January or March, and thus were not included in this analysis.

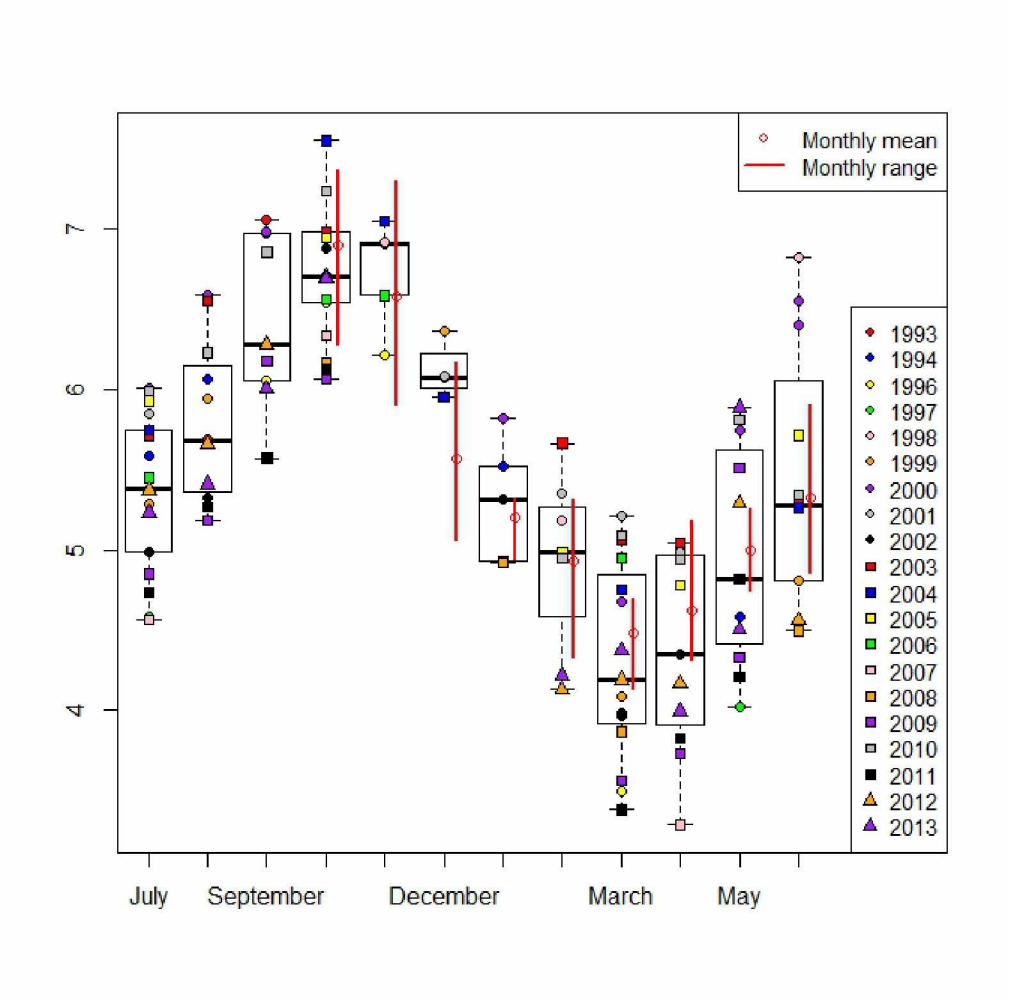


Figure S2.3-2. Boxplot of temperatures observed at a depth of 150 m at NPS Oceanographic Survey Station 4 for all years. Individual years are shown by different symbols. Monthly mean (red diamond) and range (red line) of temperatures observed at the stationary mooring near Station 4 between October 2013 and June 2014.

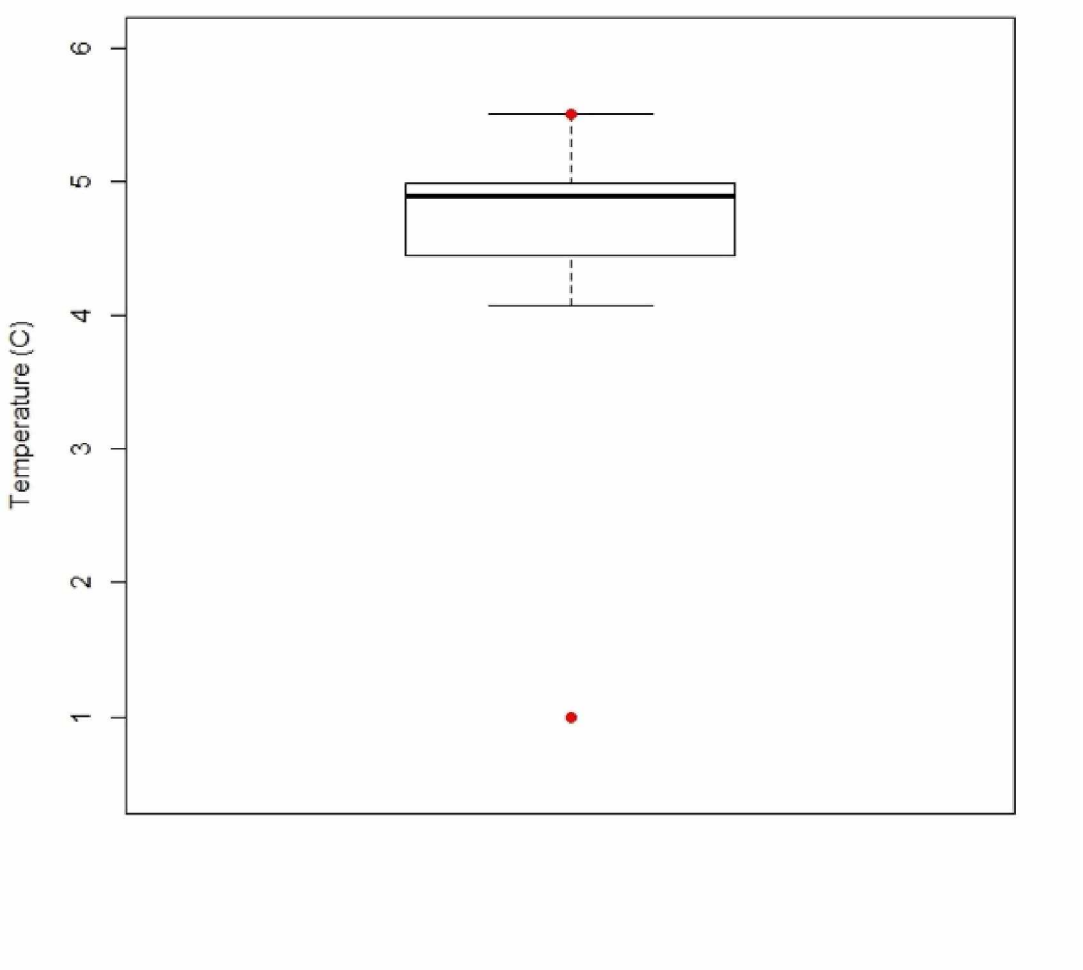


Figure S2.3-3. Boxplot of maximum temperatures for depths > 100 m at all NPS oceanographic sampling stations within the MPA sampled during January or February of "cold years" (2005 – 2014).

2.9.4 IACUC Approval letter

(907) 474-7900
(907) 474-5639 fax
fyiacuc@uaf.edu
www.uaf.edu/iacuc

UAF UNIVERSITY OF ALASKA FAIRBANKS

Institutional Animal Care and Use Committee

909 N Koyukuk Dr. Suite 212, P.O. Box 757270, Fairbanks, Alaska 99775-7270

March 24, 2011

To:	Andrew Seitz Principal Investigator
From:	University of Alaska Fairbanks IACUC
Re:	[215255-2] Multiscale movement of demersal fishes in Alaska

The IACUC reviewed and approved the New Project referenced below by Designated Member Review.

Received:	March 16, 2011
Approval Date:	March 24, 2011
Initial Approval Date:	March 24, 2011
Expiration Date:	March 24, 2012

This action is included on the March 25, 2011 IACUC Agenda.

The PI is responsible for acquiring and maintaining all necessary permits and permissions prior to beginning work on this protocol. Failure to obtain or maintain valid permits is considered a violation of an IACUC protocol, and could result in revocation of IACUC approval.

Chapter 3. Effect of study area bathymetric heterogeneity on parameterization and performance of a depth-based geolocation model for demersal fishes[1]

3.1 Abstract

State-space geolocation models can provide valuable information on the large-scale movements of many fish species. The sensitivity of such complex models to model assumptions and fixed parameters is rarely assessed quantitatively, yet is important for interpretation of results and adaptation for new species or different geographic regions. We hypothesized that parameterization and performance of a discrete Hidden Markov Model (HMM) with a Gaussian depth-based data likelihood for demersal fishes first implemented in the flat terrain of the North Sea would be affected by the more heterogeneous depths found in the North Pacific Ocean. We ran the HMM on depth data from simulated random walk movement trajectories in flat, sloping, and heterogeneous study areas in the North Pacific Ocean where known depth distributions in each model grid cell were provided by high-resolution (5 m) multibeam bathymetry. Performance was compared among different data likelihood specifications and grid sizes in each area. In the flat study area, depth distributions for all grid sizes were normal or uniform and performance was similar across grid sizes. In contrast, cell depth distributions in the sloping and heterogeneous study areas were normal or uniform only in the smallest grid sizes and performance decreased with increasing grid size. Overall model performance was highest in the heterogeneous and sloping areas for small grid sizes and in the flat area for large grid sizes. A new method for specifying grid cell depth variance based on study area slope performed better than obtaining variance from adjacent grid cell values for larger grid sizes in heterogeneous and sloping areas. The estimated value of diffusion was also sensitive to bathymetric heterogeneity and variance-specification method. These results suggest that the degree of study area heterogeneity should be considered when choosing fixed parameters such as likelihood and grid size, and when interpreting the model results. In addition, this approach demonstrates the need for sensitivity analyses when using the model on a new species or in a new study area.

[1] Nielsen, J.K., F. Mueter, M. Adkison, T. Loher, S. McDermott, and A.C. Seitz. In review, Ecological Modeling.

3.2 Introduction

Demersal fishes are a major component of commercial fisheries in the North Pacific Ocean, particularly in Alaska, USA. In 2015, 2.2 million tons of groundfishes comprised 80% of the total commercial harvest and 52% of the total ex-vessel revenue from commercial fisheries off Alaska (Fissel et al., 2016). Demersal fish species such as Pacific halibut *Hippoglossus stenolepis,* Pacific cod *Gadus macrocephalus*, and sablefish *Anoplopoma fimbria* are known to undertake large-scale ontogenetic and seasonal spawning migrations (Shimada and Kimura, 1994; Loher and Seitz, 2006; Webster et al., 2013; Hanselman et al., 2014) during which fish may move among management areas or cross international boundaries. Therefore, understanding migration characteristics such as timing, pathways, and spawning locations is important for the management of these demersal fish species.

Detailed information on fish migrations can be provided by electronic tags that continuously record data on environmental variables such as depth and temperature while attached to fishes at liberty. Data collected by the tags are matched to maps of environmental variables in the study area to determine likely locations of the tagged fish between release and recovery, a process known as geolocation. Geolocation methods vary based on fish behavior, tag type, and geographical region. For example, light intensity and sea surface temperature (SST) data are commonly used for estimating position of pelagic fishes such as tuna (Scombridae) (Teo et al., 2004). However, methods for geolocation of demersal fish species rely primarily on depth data (Hunter et al., 2004; Pedersen et al., 2008), but may include light-based longitude if fish occupy depths < 150 m (Seitz et al., 2006), as well as bottom temperature (Le Bris et al., 2013; Liu et al., 2017) or salinity information (Neuenfeldt et al., 2007).

Geolocation models estimate plausible movement paths of fishes by combining tag data with knowledge about how animals are likely to move through the study area (e.g., swim speed or movement pattern). Most recently-developed geolocation models are state-space models (SSMs) which consist of 1) a fish movement model that predicts possible locations of the tagged fish at the next time step, and 2) a data likelihood model that matches data from the tagged fish to features of the study area at each time step. SSMs are ideal for geolocation because they can accommodate noisy data as well as quantify uncertainty

in fish location estimates and other movement parameters such as movement speed (Patterson et al., 2008). SSMs for geolocation can either be parametric, such as a Kalman filter, or non-parametric, such as a Hidden Markov Model (HMM) or a particle filter. Parametric SSMs are useful for geolocation of open-ocean pelagic fishes, which are characterized by Gaussian data types (light) and unconstrained behavior (Nielsen et al., 2006). Non-linear SSMs are more suitable for demersal fishes, which are characterized by non-linear data types, such as depth, and potentially constrained near-shore behavior (Andersen et al., 2007; Pedersen et al., 2008).

An HMM developed for geolocation of Atlantic cod in the North Sea (Pedersen et al., 2008; Thygesen et al., 2009) has provided a flexible model framework that can be readily adapted for geolocation of different fish species in different regions (Le Bris et al., 2013; Neat et al., 2014; Neilson et al., 2014; Woillez et al., 2016; Biais et al., 2017; Liu et al., 2017; Strøm et al., 2017). Adapting this HMM for different applications primarily involves changes to the data likelihood model (i.e., the specific way in which tag data are matched to maps of environmental variables in the study area) and selection of the appropriate model grid size. However, adapting the HMM for different geolocation applications is not a trivial process because both data likelihood and grid-size specifications depend on characteristics of the study area, tag data, and fish behavior. For example, data likelihood models depend on the type of geolocation data collected by the tag, tag measurement resolution and frequency, strength of environmental gradients in the study area, and accuracy and resolution of available maps of environmental characteristics. An understanding of fish behavior is also important for the specification of data likelihood (e.g., whether the fish is expected to be at the surface, at the bottom, or moving throughout the water column). The optimal model grid size depends on the spatial gradients of geolocation variables in the study area as well as the speed at which fish are expected to move through it.

To adapt the Pedersen et al. (2008) HMM for the geolocation of demersal fishes in the North Pacific Ocean, we chose a data likelihood model that is based on matching the maximum depth of the fish at each time step to bathymetric maps of the study area. For this likelihood method, demersal fishes are assumed to be on or near the bottom at some point during each time step. The maximum-depth likelihood

method was chosen because bathymetry gradients can be steep in the North Pacific Ocean, and bathymetric map resolution and accuracy are improving steadily in many areas (Zimmermann and Prescott, 2015) as additional data are collected. In contrast, gradients in tidal amplitude and phase required for the tidal geolocation method that has been utilized for geolocation of demersal fishes in other geographic regions (Hunter et al., 2004; Pedersen et al., 2008) are weak in many areas of the North Pacific Ocean (Danielson et al., 2011). In addition, because the maximum-depth likelihood can be used for both detailed (e.g., physically-recovered) and summarized data from Pop-up Satellite Archival Tags (e.g., transmitted through Argos satellites), it can be applied to most demersal fish geolocation data sets. Therefore, our data likelihood model is based on maximum daily depth, but other types of geolocation data (e.g., bottom temperature, light-based longitude, or tidal amplitude and phase) can be added when available.

The use of depth records and bathymetry maps as the primary source of geolocation data has several implications for data likelihood and grid-size specification given the degree and nature of bathymetric heterogeneity present in the study area. Any method that provides information about which grid cells are most likely to be occupied by the tagged fish can be used to specify the data likelihood (Thygesen et al., 2009). In applications for pelagic fishes, depths shallower than the maximum depth that was recorded by the tag are simply discarded as possible fish locations during that time-step by assigning a zero value to cells with depths shallower than the maximum recorded depth of the fish; a value of one is assigned to cells with depths greater than the maximum depth of the fish (Woillez et al., 2016; Biais et al., 2017; Strøm et al., 2017). In contrast, maximum depth likelihoods for demersal fishes tend to assign likelihood values based on an assumed normal distribution of depths within each grid cell (Pedersen et al., 2008; Le Bris et al., 2013; Liu et al., 2017). However, depending on the spatial scale of depth variation in the study area and a given grid size, the choice of a normal distribution within each grid cell may not be optimal. In previous applications, the variance used to specify the normal probability distribution within each cell has been obtained from the depth values of adjacent grid cells (Pedersen et al., 2008; Le Bris et al., 2013; Braun et al., 2017; Liu et al., 2017). This method may not be appropriate if the grid size used

for the model is larger than the scale of autocorrelation of depths in the study area. Additionally, large grid-cell sizes may not provide the optimal resolution for detecting depth gradients in areas with highly heterogeneous depths (e.g., if all cells have similar means and variances at larger grid cell sizes). Although data likelihoods and grid sizes are reported in previous studies that use HMM for geolocation (e.g., Le Bris et al., 2013; Woillez et al., 2016), sensitivity analyses that explore the effect on model accuracy and precision of choosing different options for data likelihood or grid size are rarely reported (but see Braun et al., 2017, Appendix S1). Such analyses are important for selecting optimal grid sizes and likelihoods as well as for identifying potential strengths and weaknesses of the model given the degree of habitat heterogeneity in the study area. In addition, such analyses provide a measure of confidence in model outputs such as uncertainty estimates.

Because the data likelihood model for demersal fishes in the North Pacific Ocean is based primarily on maximum depth, and the degree of depth heterogeneity varies by region, undertaking such sensitivity analyses is important for selection of optimal grid size and likelihood specifications in a given study area. Therefore, we perform a series of sensitivity analyses to understand the effect of grid size and data likelihood specifications on model performance in three regions of Alaska that differ in degree of bathymetric heterogeneity. In each area, we apply the HMM to simulated trajectories of fish movement and compare performance (estimated location and location error) for different likelihood and grid size specifications. We present a new method for determining grid-cell depth variance used in likelihood specification. We anticipate that the sensitivity analysis procedures performed here will serve as a guideline for researchers planning studies of demersal fish movement in other regions based on the degree of depth heterogeneity present in their study area or studies that involve other gridded geolocation variables such as SST.

3.3 Materials and Methods

3.3.1 Model description

The HMM geolocation model (Pedersen et al., 2008) consists of a general model framework that is modified for different applications through specification of the data likelihood model and selection of

the appropriate grid size. The HMM geolocation model operates in discrete time and space, where the study area is divided into discrete grid cells. The data likelihood model specifies the way data measured by the tag are related to environmental conditions present in each grid cell at each time step. The underlying movement model is a random walk applied to grid cells as a two-dimensional symmetrical diffusion kernel. The model can incorporate multiple behavioral states by allowing the size of the diffusion kernel to vary for different types of behavior (e.g., foraging vs. migrating). Behavioral states are not estimated by the model, but are specified prior to modeling based on auxiliary analyses that vary depending on the species and type of data available. This approach greatly simplifies the model and preserves generality for application to different species or tag types. However, the model can estimate the size of the diffusion kernel used for each movement state through maximum likelihood.

The data likelihood model chosen for geolocation of demersal fishes in the North Pacific Ocean is based on the maximum depth likelihood. The likelihood of observing the maximum recorded depth of the fish at each time step given the depth distribution in each cell in the study area is obtained by integrating the probability distribution of depths in each cell between the limits of the tag maximum depth value +/- the tag measurement uncertainty (Le Bris et al., 2013). Likelihoods for different geolocation variables are combined by cell-wise multiplication at each time step, as long as they are not correlated (Le Bris et al., 2013; Liu et al., 2017). In the event of missing data, all grid cells for the likelihood of that geolocation variable are assigned a value of one for that time step.

The model is executed through a "forward filter" that iteratively 1) advances the probability of the fish location in the study area at the next time step by convolution with an isotropic diffusion kernel (movement update), and 2) updates the cell probabilities by element-wise multiplication with the data likelihood model for the next time step (data update). Once the last time step is reached, a backward smoothing step is performed that adjusts probability estimates obtained from the forward filter to reflect knowledge of the end location. The smoothed probability estimates can be viewed directly for each time step or they can be summarized over certain time periods by calculating a residence distribution.

Movement pathways are reconstructed by global decoding with the Viterbi algorithm (Viterbi, 1967) to produce the most probable track (Pedersen et al., 2008; Thygesen et al., 2009; Woillez et al., 2016; Liu et al., 2017; Strøm et al., 2017) or by local decoding based on the maximum, mean, or mode of the smoothed probability distribution at each time step (Thygesen et al., 2009; Woillez et al., 2016; Biais et al., 2017). Local decoding methods for path reconstruction, such as the weighted mean of the smoothed probability at each time step, provide an estimate of uncertainty in the form of 95% probability ellipses calculated from the weighted covariance at each time step (Biais et al., 2017; Doherty et al., 2017).

To determine the sensitivity of the model to grid size choice, likelihood specification, and study area depth heterogeneity, we assessed model performance on simulated movement paths in three regions of the North Pacific Ocean that vary in degree of depth heterogeneity (Figure 3.1). In each region, we simulated movement paths within 6 x 12 km study areas where high-resolution (5m) multibeam sonar bathymetry sets were available. High resolution bathymetry data sets provided a known distribution of depths for each grid cell size and allowed accurate depths to be extracted from simulated trajectories. In the Norton Sound study area located in northwestern Alaska (Figure 3.1A), large expanses of relatively flat terrain are present and depths range from 52 to 61 m (NOAA NCEI, 2004). This study area is referred to as "flat" in the analyses. In the Central Gulf of Alaska study area (Figure 3.1B), the bathymetry is primarily sloping and depths range from 48 m to 179 m (National Ocean Service Hydrographic Surveys, 2004). We refer to this area as "sloping" in the analyses. The Glacier Bay study area, located in the eastern Gulf of Alaska (Figure 3.1C), features heterogeneous glacial topography. Depths in this study area range from 13 m to 340 m (Carlson et al., 2002). We refer to this area as "heterogeneous" in the analyses. In each study area, the HMM was applied to the depths extracted from simulated trajectories with different combinations of grid sizes and likelihood specifications and estimated locations were then compared to known (simulated) locations.

3.3.2 Study area bathymetric characteristics

For each study area, we characterized depth gradients and depth distribution characteristics for different grid cell sizes. First, we aggregated the multibeam sonar bathymetry (5-m resolution) data into a

series of larger grid sizes by calculating the mean of the 5-m bathymetry values within each larger grid size. Then we calculated slope for each grid size using the aggregated bathymetry grids with the R package 'SDMTools' (VanDerWal et al., 2014). To characterize properties of statistical depth distributions in aggregated grid cells (e.g., all 5-m grid-cell depth values within each larger grid size), we determined skewness (a measure of asymmetry) and kurtosis (a measure of tail thickness) values using the R package 'fitdistrplus' (Delignette-Muller and Dutang, 2015). We plotted slope, skewness, and kurtosis values for all grid sizes in each area as empirical cumulative frequency distributions. To allow comparison with typical values for uniform, normal, logistic, lognormal, exponential, gamma, and beta distributions, we summarized skewness and kurtosis values in Cullen and Frey plots (Cullen and Frey, 1999). For comparison, we also visualized the observed distribution of depths across all cells in a given study area for each grid size.

3.3.3 HMM performance comparisons

3.3.3.1 Simulated trajectories

To assess model performance under different conditions (study area depth heterogeneity, likelihood specification, and model grid size), we generated simulated fish-movement pathways that allowed a direct comparison between estimated and known (simulated) locations in each study area. We chose a random walk (equivalent to diffusion) as the underlying movement pattern for the simulations to match the movement model used in the HMM. Trajectories were simulated using the formula:

$x_t = x_{t-1} + \varepsilon_{xt}$, $y_t = y_{t-1} + \varepsilon_{yt}$, and $\varepsilon_t = N(0, \sigma^2)$

where x_t and y_t are the X and Y coordinates at time t, x_{t-1} and y_{t-1} are the X and Y coordinates at the previous time step, and ε_t is normally distributed error with a mean of 0 and standard deviation of σ.

Trajectories were simulated using σ = 200 m, which corresponded to a diffusion coefficient (D) of 0.02 km^2 per time step. At each time step, the new location was compared to bathymetry to ensure that simulated tracks did not fall on land and a new error was simulated otherwise. Each simulation consisted of 200 steps. We extracted depth at each time step from high-resolution (5 m) bathymetry and used those

extracted depths to simulate data that would be recorded by an archival tag attached to the simulated fish. Tidal stage data specific to each study area were obtained from the nearest NOAA water level station (available at tidesandcurrents.noaa.gov) and added at 30-minute intervals to extracted depths in the simulated movement trajectories. Random uniform error between the interval of typical archival-tag measurement resolution (e.g., ±1 m) was added to the depth record to simulate noise due to tag-measurement resolution. The result was a simulated (known) movement trajectory accompanied by a simulated set of tag-depth data.

3.3.3.2 Likelihood and grid size treatments

Maximum depth likelihood values were obtained by integrating the probability density of depths in each grid cell by the limits of the simulated tag maximum depth plus and minus the tag measurement uncertainty (Le Bris et al., 2013). Conceptually, this is the intersection of a uniform probability distribution representing the measurement uncertainty of the tag with the probability distribution of depths in each cell (Figure 3.2). This approach allowed incorporation of measurement uncertainty due to the tag measurement resolution and change in depth due to tides. Decisions needed to specify this likelihood were 1) the type of statistical distribution used to define the probability density function for depths in the study area (e.g., normal vs. log-normal), 2) methods for estimating parameters for each statistical distribution (e.g., variance of a normal distribution), and 3) the value of tag measurement uncertainty used to integrate the probability distribution.

We conducted preliminary analyses to narrow the number of potential treatments to those for which model performance was the most sensitive. Because model performance with statistical distributions other than the normal distribution did not have an appreciable effect on performance in the preliminary analysis, we used the normal distribution for all treatments. However, preliminary analyses did suggest that the model was sensitive to the method used to determine the variance of the normal distribution and the value used for tag-measurement uncertainty. Therefore, our likelihood treatments consisted of three methods for specifying the variance of the normal distribution used to describe the statistical distribution of depths in each cell. First, because our bathymetry data sets had a higher

resolution than the model grid sizes, we employed a within-cell method that consisted of the standard deviation of high-resolution depth values used to form each aggregated grid cell. Second, the "roughness" method consisted of calculating the standard deviation of the depths in the eight adjacent grid cells. The roughness method has been utilized to specify the variance in previous HMM applications where the bathymetry resolution is the same as the model grid size (Le Bris et al., 2013; Liu et al., 2017). Finally, as an alternative to the roughness method, we developed a "slope" method, where variance in each cell was estimated based on relating the observed slope to observed within-cell variances using a Generalized Additive Model (GAM) in areas where high-resolution multibeam bathymetry data were available. Detailed descriptions of the different likelihood treatments are provided in Supplementary Data, Appendix A. We compared model results using within-cell, roughness, and slope methods for determining variance in the heterogeneous and sloping study areas. In the flat study area, we compared only the within-cell and roughness methods, as slopes were negligible in this area.

We also tested three values of depth uncertainty for each region. Each study area has a known depth uncertainty that is a combination of tag depth resolution (assumed to be 1 m in simulations) and the tidal range used for simulations (flat = 1 m, sloping = 2 m, heterogeneous = 4 m). We tested a "low" value (half the magnitude of the known uncertainty), a "true" value (known uncertainty), and a "high" value (twice the magnitude of the known uncertainty). To understand the effects of grid size on model performance, we ran the model with a progression of grid sizes (100, 200, 400, 1000, and 2,000 m) in each study area.

3.3.3.3 HMM estimation

We ran the HMM on the extracted depths from simulated trajectories (n=100 per area) using different combinations of data likelihoods and bathymetry grid sizes. Data likelihoods were produced in R and sent to Matlab to run the model (forward filter, backward smoothing, and MPT computation) with the HMM toolbox code provided by Martin Pedersen (DTU, Denmark). Each HMM run produced a 3-D array of smoothed probability estimates (e.g., spatial probability distribution of the simulated trajectory location at each time step), a residency distribution that summarized the estimated locations over all time

steps, an estimate of diffusion (and uncertainty in the diffusion estimate), probability polygons for each time step, and reconstructed movement paths from both the Viterbi algorithm (MPT) and the weighted mean methods. An example of a simulated trajectory in the heterogeneous study area processed with the same likelihood treatment at 200 m and 1000 m grid cell sizes demonstrates resulting reconstructed pathways (Figure 3.3A) and residence distributions (Figure 3.3B). An upper limit of twice the value of diffusion used to simulate trajectories (0.04 km^2) was imposed on the simulations to prevent unrealistic diffusion estimates.

3.3.3.4 Assessing HMM performance

To assess the performance of different likelihood/grid size combinations, we calculated four metrics derived from HMM estimates and known parameters from the simulations. Performance metrics assessed accuracy of reconstructed pathways, depth bias for reconstructed pathways, whether or not probability polygons at each time step contained the simulated trajectory location (coverage probability), and the accuracy of model estimates for the diffusion coefficient (D).

3.3.3.4.1 Mean Absolute Error of reconstructed movement paths

We assessed accuracy of reconstructed movement paths using a mean absolute error (MAE) performance metric. This metric calculates the mean absolute distance between the known locations from simulated trajectories and positions estimated by the HMM at each time step (n = 200 per trajectory). We calculated MAE for both the MPT (most probable track) and the weighted mean methods of obtaining location estimates. Because HMM estimated locations are individual grid cells, we used the center of each grid cell to measure the distance between simulated and estimated tracks.

3.3.3.4.2 Depth bias

To determine whether different likelihood and grid-size treatments resulted in a systemic bias toward deeper or shallower depths for reconstructed movement paths, we compared the depth from estimated pathways to depth at simulated trajectory locations. At each time step, the depth at the simulated trajectory location was subtracted from the depth at the model-estimated location using the 5-m resolution multibeam bathymetry and the average difference over all time steps was calculated. The bias

metric was calculated for both the MPT and the weighted mean methods of path reconstruction. Negative values correspond to bias toward shallower depths. Results from the bias performance metric were tested for significant difference from zero using a two-tailed Wilcoxon rank sums test. For treatments found to be significantly different from zero, the direction of the bias was recorded.

3.3.3.4.3 Coverage probability

To determine whether probability quantiles at each time step were either over- or underestimating actual error, we assessed the coverage probability at the 80% and 99% probability quantiles. We constructed polygons around the smallest number of grid cells that together encompassed 80% and 99% of the probability at each time step and determined whether or not the known location at that time step was within each polygon. The proportion of locations that were inside their respective error polygons at each time step were reported as the coverage probability metric.

To determine whether coverage probability at large grid sizes could be improved in the heterogeneous areas, we applied a range of additional likelihood treatments to the 1000 m grid size. Additional treatments consisted of different combinations of statistical distribution types for bathymetry (normal, lognormal, T distribution), adding extra variance to the bathymetry probability distribution, switching from a uniform to a Gaussian distribution for tag measurement error, and using a fixed value of diffusion. Boxplots of additional treatment results were produced for 80% and 99% coverage probability metrics.

3.3.3.4.4 Diffusion

We compared the value of diffusion estimated by the HMM for different treatments to the value used to simulate fish movement trajectories. Because simulating pathways in a bounded area could result in a different effective value of D than was used to generate the simulated paths, we calculated an empirical value to use for comparison with values estimated by the HMM. The empirical value for D was determined by calculating the mean squared displacement (MSD) from the initial and final locations of simulated paths for all three areas combined (n = 300) using the formula MSD = 4Dt (Börger and Fryxell, 2012). Results from the diffusion performance metric were tested for significant difference from the

empirical value of diffusion for simulated trajectories using a two-tailed Wilcoxon rank sums test. For treatments found to be significantly different, the direction (e.g., higher or lower) was recorded.

3.3.3.4.5 Summarizing performance

To summarize results among treatments by grid size, treatments were ranked according to their performance for each metric. The best performance was defined for each metric as 1) MAE: lowest values, 2) bias: closest to zero, 3) coverage probability: closest to 80% or 99%, and 4) treatment-estimated diffusion: closest to the empirical value of diffusion. Treatments were ranked by calculating decile values for each treatment, ranking all treatments by decile, then averaging rank scores from all deciles for each treatment. For the bias metric, only the median was used to rank treatments because the full range of bias measurements reflects the degree of depth heterogeneity in each study area. All statistics and analyses were computed using the R package (R Core Team, 2017).

3.4 Results

3.4.1 Study area characteristics

Study area slope distributions differed strongly among the three study areas. Slope values were greater in the heterogeneous area compared to the sloping area by a factor of 10, and values in the sloping area were greater than the flat area by a factor of 10 (Figure 3.4). In the heterogeneous area, slopes were largest in the 100-m grid size but decreased as grid size increased. In the sloping area, 100-m – 400-m grid sizes had similar slopes, but slope decreased for 1000-m and 2000-m grid sizes. In the flat area, slopes were very small for all grid sizes.

Differences in kurtosis and skewness values for depth distributions in each grid cell were also apparent among study areas. In the heterogeneous area, kurtosis values increased with grid cell size. Values for the 1000-m and 2000-m grid sizes were similar and larger than 100-m – 400-m grid sizes. Kurtosis values in the sloping area were slightly smaller than the heterogeneous area and values for the 1000-m grid were more similar to values for the 400-m grid. Negative skew (i.e., skew toward shallower depths) was observed for the 1000-m and 2000-m grid sizes in the heterogeneous areas and the 2000 m

grid size in the sloping area. In the flat area, both kurtosis and skewness values were low and similar among all grid sizes.

When compared with various types of distributions in the Cullen Frey plots, small grid-cell sizes in all study areas were concentrated in the vicinity characteristic of uniform to normal (Figure 3.5). As grid-cell size increased, distributions in the heterogeneous (Figure 3.5A) and sloping (Figure 3.5B) areas became more consistent with non-normal and non-uniform distributions. However, distributions in the flat area were similar across all grid-cell sizes and consistent with normal or uniform distributions (Figure 3.5C).

3.4.2 Assessing HMM performance

3.4.2.1 Mean Absolute Error of reconstructed movement paths

In general, Mean Absolute Error (MAE) for the weighted mean method of path reconstruction increased as grid size increased for the heterogeneous and sloping areas, but was consistent across grid size for the flat area (Figure 3.6A). Model performance was best (smallest MAE values) in the heterogeneous area, intermediate in the sloping area, and lowest in the flat area for the 100-m – 200-m grid sizes. Within each grid size, treatment performance varied most by variance method in the heterogeneous and sloping areas, where the roughness method consistently had lower performance compared to within-cell and slope methods. In the sloping area, the within-cell and slope methods had similar performance, whereas the within-cell method consistently performed better in the heterogeneous area. In the flat area, treatment performance varied most by depth uncertainty, where the true value performed best, the high value was intermediate, and the low value performed the worst. However, the within-cell method of variance outperformed the roughness method for the true and high values of depth uncertainty.

In general, trends for the MPT method of reconstruction were similar to the weighted mean method (Figure 3.6B). However, the weighted mean performed better than the MPT method in the 100-m – 200-m grid sizes in the heterogeneous and sloping areas. For MPT reconstruction, the roughness variance method performed better than the within-cell and slope methods in the smallest grid sizes in the

heterogeneous and sloping areas. Overall, best results for both the weighted mean and MPT methods of path reconstruction were observed for the heterogeneous area at small grid sizes and the flat area at the largest grid sizes (Figure 3.7A and 3.7B).

3.4.2.2 Depth bias

In general, the median values for depth bias (e.g., the mean difference between depth at simulated locations and the depth at model-estimated locations) were small for both the weighted mean and MPT methods of path reconstruction (Table 3.1 A and B). The greatest biases were observed in the heterogeneous area in the direction of deeper water for grid sizes 400 m – 1000 m. The roughness method of variance was not biased in the 1000-m – 2000-m grid sizes for either the weighted mean or the MPT paths. Some bias in the direction of shallower water was observed in the smaller grid sizes in the heterogeneous area. Although many treatments in the sloping and flat areas were observed to have significant bias, the magnitude of the difference from zero consisted of only a few cm. Although median bias values were similar between the weighted mean and MPT methods of path reconstruction, the range in MPT bias tended to be much smaller than the range in weighted mean bias values (Figure S3.2-1). In general, bias was lowest in the flat study area for both the weighted mean and MPT path reconstruction methods (Figure 3.7C and 3.7D).

3.4.2.3 Coverage Probability

In general, coverage probability at nominal levels (80% and 99% error polygons) was only achieved for the 99% error polygons at the smallest grid sizes (100-m – 200-m) in the heterogeneous study area (Figure 3.8). Coverage probability approached nominal levels for the 99% error polygon for some treatments at the 400 m grid size in the heterogeneous area, at grid sizes 100 m – 400 m in the sloping area, and at grid sizes 100 m – 1000 m in the flat study area. The true and high levels of depth uncertainty outperformed the low level of depth uncertainty for grid sizes 100 m – 400 m in all areas and the 1000-m – 2000-m grid size in the flat area, but performance did not vary by depth uncertainty in the 1000- m – 2000-m grid sizes in the heterogeneous and sloping areas. In general, the within-cell and slope variance methods outperformed the roughness method for grid sizes 400 m – 2000 m in the heterogeneous

and sloping areas. The slope method performed as well or better than the within-cell method for the 2000-m grid size in both heterogeneous and sloping areas. Performance trends by treatment for 80% error polygons were the same as for the 99% error polygons but the magnitude of coverage probability was substantially lower than nominal levels compared to performance for the 99% error polygons.

Of the additional treatments performed to achieve increased coverage probability in a large (1000 m) grid cell size in the heterogeneous area, the best result was achieved by introducing a Gaussian tag measurement uncertainty distribution with a s.d. of 5 m, a Gaussian bathymetry grid-cell distribution with an extra 5 m of variance, and a fixed diffusion coefficient of 0.04 km^2 (Table 3.2, Figure S3.2-2). However, the best-performing treatment only increased the median 80% coverage probability from 52% to 60%. Fixing the value of diffusion at 0.012 km^2 (the empirical value of diffusion for simulated trajectories) resulted in substantial improvements in MAE but not coverage probability. In general, treatments that increased MAE resulted in decreased performance for coverage probability. Adding extra variance to the grid cell variance had little effect on performance by itself, but when combined with use of the Gaussian tag-measurement error, performance was increased. Specifying other types of distributions (lognormal, T distribution with 1 and 2 degrees of freedom) for the grid cell probability density function resulted in reduced coverage probability compared to Gaussian distribution. Overall, performance was highest in the heterogeneous area for the smallest grid sizes (100 m and 200 m), the sloping area for the 400 m grid size, and the flat study area for the largest grid sizes (1000 m and 2000 m (Figure 3.7E and 3.7F).

3.4.2.4 Diffusion

In general, diffusion was overestimated for all grid sizes in the heterogeneous area and larger grid cells (400 – 2000 m) of the sloping area (Table 3.1C, Figure 3.9). Performance was best for the 100-m – 200-m grid sizes in the sloping area, where the high level of uncertainty performed better than the true and the low. In the flat area, performance differed strongly by level of depth uncertainty. Diffusion estimates for the low treatment consistently approached the diffusion threshold value of 0.04 km^2, while estimates were closest to actual values for the true and high treatments. In contrast, differences in

performance in the heterogeneous area were mostly due to variance method, where the roughness method consistently provided estimates that were closer to the actual value of diffusion compared to the within-cell or slope methods. Overall optimal performance was observed in the heterogeneous and sloping areas at small grid sizes and in the flat area at larger grid sizes (Figure 3.7G).

3.5 Discussion

The HMM geolocation model developed by Pedersen et al. (2008) has provided a robust framework that can be adapted for different geolocation applications (e.g., study areas or fish species). When adapting the HMM for specific applications, it is necessary to 1) choose an appropriate data likelihood model, 2) select a grid size, 3) decide whether to estimate diffusion with the model or to use prior knowledge to designate the size of the diffusion kernel. The results of our simulation exercises suggest that the degree of bathymetric heterogeneity in the study area can impact the model results and should be considered when making these decisions. In addition, our work provides several implications for interpretation of results given grid size, likelihood specification methods and degree of habitat heterogeneity.

3.5.1 Data likelihood model parameterization

3.5.1.1 Variance specification

The choice of method used to specify grid cell variance affected model performance for large grid sizes in heterogeneous and sloping areas, where grid-cell depth distributions departed from normal. This is an important finding because the roughness method, which performed worse than the other two methods, has frequently been employed in other HMM geolocation studies (Le Bris et al., 2013; Braun et al., 2017; Liu et al., 2017). Our research suggests that in areas with heterogeneous topography, different methods for obtaining variance should be explored and the method that provides the most realistic values should be chosen.

In the current study, the method of obtaining grid-cell variance by linking known grid-cell variance to grid-cell bathymetric slope in areas with high-resolution multibeam data provided better results than the roughness approach for larger grid sizes in heterogeneous and sloping areas. Performance

with the slope method was sometimes equivalent and sometimes better than the within-cell method. Therefore, this study has provided a potential alternative method for specifying variance based on auxiliary data. However, additional verification is required to ensure the validity of applying the method outside of areas with high resolution depth data (the multibeam areas) that were used to determine the relationship between s.d. and grid-cell slope. For example, the relationship between cell variance and slope could be determined by one multibeam area and tested in another multibeam area within the same larger geographic region.

The approach of using auxiliary data to define variance may be helpful for other gridded geolocation variables aside from depth. For example, some SST data sets such as multi-scale ultra-high resolution (MUR) SST are accompanied by estimates of error in each grid cell; these should be investigated for use in specifying cell variance rather than, or perhaps in addition to, utilizing the roughness method in heterogeneous areas. Using an Expectation-Maximization framework to fit the model (Woillez et al., 2016), where variance parameters are estimated rather than specified *a priori*, may also be an option for using the HMM in heterogeneous areas.

Although the roughness method of determining variance decreased performance at larger grid sizes for the MAE metric, it did have advantages in other situations. Because it performed best in the smallest grid sizes, where variance values tended to be larger than the other two methods (Supplementary Data Figure A-2), adding extra variance to the within-cell variance may increase performance at smaller grid sizes. The roughness method also performed better in the diffusion coefficient estimation and bias performance metrics, which suggests that increasing the grid cell variance may also improve performance in these areas. Improved performance of the coverage probability metric for treatments with extra added variance (Table 3.2) supports this observation.

3.5.1.2 Depth measurement uncertainty

Determining the likelihood value by integrating between the limits of tag measurement resolution (Le Bris et al., 2013) has been featured in several recent applications of the HMM (Braun et al., 2017; Liu et al., 2017; Braun et al., 2018). This approach is valuable in that it allows explicit incorporation of tag

measurement uncertainty into the calculation of the likelihood. However, our research demonstrates that the performance of the model can be sensitive to the magnitude of the value chosen. In general, all-around performance was best when the chosen value reflected the true level of uncertainty associated with tag measurement. Although the low value resulted in slight increases in performance for the MAE metric in the 100-m – 400-m grid sizes of the heterogeneous and sloping areas, performance for the coverage probability and diffusion metrics was worse for the low value compared to the true and high values. For larger grid cells, our results indicate that it is better to overestimate than underestimate the depth measurement uncertainty. This may be more important for flat areas, where differences in performance by depth uncertainty level were more pronounced than the heterogeneous and sloping areas.

The use of a Gaussian distribution to represent tag measurement uncertainty may represent an improvement in the model for our application. Tag measurement resolution may be uniform for some types of tag sensors, but with added tidal data, the depth measurement uncertainty becomes Gaussian. Demersal fish data likelihood models that include tidal geolocation (Pedersen et al., 2008) explicitly account for change in depth due to tides. In contrast, our data likelihood model does not include tidal data, and therefore we must account for uncertainty in tag measurement due to tidal fluctuations in another way. In areas of the North Pacific Ocean, the tidal amplitude can exceed 3 m depending on geographic location (Danielson et al., 2011). For large-scale demersal fish applications in this region, it may be desirable to use a Gaussian depth uncertainty distribution that is wide enough to incorporate potential movement of the fish to areas with a greater range of tides than are found at the starting location.

3.5.2 Grid size selection

Selection of grid size is a key consideration for implementation of the HMM. The resolution of available maps, the size of the study area in which the animal moves, the strength of geolocation gradients in the study area, the movement speed of the fish, and spatial resolution required by research goals (e.g., residence time in Marine Protected Areas) should all be considered when selecting a grid size. In addition, when the data likelihood model combines multiple environmental variables (e.g., both depth and bottom temperature), the choice of grid size is constrained to the lowest grid resolution available for each variable.

Increases in time and computer power needed for processing are also a consideration for grid size selection, but with modern computing power this is becoming less of a concern.

Our research suggests that, in addition to these other considerations, the statistical distribution of depths in each cell size plays a role in grid-size selection. Grid cells with an approximately normal distribution provided a range of cell sizes where model performance was similar within each area. Optimal statistical distribution grid sizes were 100 m – 200 m in the heterogeneous area, 100 m – 400 m in the sloping area, and all grid sizes in the flat area. Reduced performance of the 2000-m grid in the flat area despite an approximately normal depth distribution within grid cells indicates that other factors such as movement rate relative to grid size in low-gradient areas may play a role in performance. Our range of grid sizes tested did not include grid sizes substantially smaller than the movement rate used to simulate the trajectories, so insights into whether this could adversely affect performance were not provided by this simulation exercise.

3.5.3 Diffusion

The decision to estimate the diffusion coefficient (D) in the model or to use a pre-determined value is influenced by study-area gradients, grid size, and likelihood-specification methods, as well as whether prior information on fish movement rates is available. Some HMM applications estimate D in the model (Le Bris et al., 2013; Woillez et al., 2016) while others pick a value based on known animal movement speeds (Braun et al., 2017; Liu et al., 2017). The model is known to be sensitive to the value of the diffusion coefficient (Woillez et al., 2016), so making this decision is a critical aspect of adapting the model for different applications.

The results of our simulations suggest that estimating diffusion in large grid cells in heterogeneous and sloping areas is associated with decreased model performance that may be compounded by the method used to specify likelihood. The model overestimated diffusion in these situations, frequently approaching its fixed upper threshold, but optimal performance for other metrics (e.g., MAE and coverage probability) was associated with estimated values of diffusion that were closer

to the value used to create the trajectories. Therefore, if accurate information on the movement of fish is available, using a known value of D might result in better model performance in these situations.

However, for small grid sizes in heterogeneous and sloping areas, where gradients are strong and grid depth distributions are approximately normal, estimating diffusion is likely to provide reasonable results for other performance metrics as well. In this case, estimating D makes sense, and estimated values of D can then be used in areas where gradients are less strong or in heterogeneous areas where larger grid-cell sizes are necessary. If estimated in the model, the diffusion coefficient should not be taken as reflecting actual animal movement speed, but rather viewed as the engine that drives the movement model.

In contrast to heterogeneous and sloping areas, estimated diffusion values in the flat area tended to be underestimated as much as overestimated. In the absence of strong environmental gradients, the model will tend to favor more direct movement pathways connecting the release and recovery locations, so the size of the diffusion coefficient will be strongly related to net displacement. The exception to this occurred for the low value of depth-measurement uncertainty, which frequently approached the threshold value set for diffusion. This may have been caused by high ratio of added noise to study-area depth gradients (e.g., added noise was ± 1 m for tag resolution and ± 0.2 m for tide, and the median slope for the 100-m grid was 0.1 m per 100 m) and reduced intersection of depth measurements with bathymetry grid cell probability density functions. Therefore, in low-gradient areas it might be worthwhile to specify a reasonable diffusion size from prior data sources rather than risk under- or over-estimating it with the model.

3.5.4 Implications for interpreting model output

3.5.4.1 Movement path reconstruction

Our comparison of two different methods for reconstructing movement paths suggests that at small spatial scales with complete data sets, the weighted mean method performs either better or equivalently to the MPT method. Although few advantages for MPT over the weighted mean method were observed in this small-scale simulation study, the MPT may have advantages in larger-scale

applications or situations with missing data or low spatial gradients. The two types of paths are calculated using fundamentally different approaches, where the MPT is the most likely sequence of states (e.g., grid cells) and the weighted mean is derived from the 3D array of smoothed probability estimates. The MPT may diverge from the smoothed estimates in areas where gradients are weak (Pedersen et al., 2011). Locations estimated with the weighted mean method may have low probability if they happen to fall between two areas of high probability (Thygesen et al., 2009). In our study, this phenomenon was demonstrated by the much-reduced range in depth differences between the simulated and estimated locations for the MPT compared to the weighted mean method. Therefore, it may be beneficial to reconstruct movement paths with both methods in situations where geolocation data are sparse or in areas of low gradients.

3.5.4.2 Coverage probability

Understanding error associated with location estimates is an important aspect of geolocation and a primary benefit of using complicated state-space models. With the HMM, estimated locations of the tagged fish can be represented by probability quantiles (e.g., 95% and 50%) at each time step (Pedersen et al., 2008). In addition, some researchers prefer to use the daily probability estimates or summarized residence distributions as indications of fish locations rather than reconstructing movement pathways (Le Bris et al., 2013). Therefore, understanding the ways in which the probability surface actually corresponds to the area that is likely to contain the true location of the fish at a specific time is crucial for interpreting HMM results.

We found that the probability quantiles derived from the probability surface at each time step did not include the true location as often as expected from the nominal coverage probability. This suggests that the model tends to underestimate the true uncertainty in location at a given time step. This was noticeable particularly for the larger grid sizes in the heterogeneous and sloping areas. It is possible that these results are due to some aspect of the model, such as backward smoothing from the recovery location that tends to result in a direct path between release and recovery locations when geolocation gradients are not informative. It is also possible that these results are due to some artifact of the simulations, or the

small and confined nature of that study area in this study. However, we believe this should be an area for future investigation because probability contours are a common way to interpret probability surfaces.

One option for increasing coverage probability is to make changes to the data likelihood model. Our results suggest that coverage probability may be improved by using a Gaussian distribution for tag measurement uncertainty instead of a uniform distribution, adding extra variance to the bathymetry probability, or increasing the depth-measurement uncertainty. However, these changes to the likelihood resulted in only marginal improvement in coverage probability and came at a cost of decreased accuracy of the reconstructed pathways for the 1000-m grid size in the heterogeneous area. This indicates that a trade-off between coverage probability and model accuracy may exist. Simulations conducted over larger scales in space and time may be needed to more fully understand potential trade-offs for coverage probability and accuracy at scales relevant to fish migration.

Coverage probability may also be increased by increasing the diffusion coefficient (D). A previous study that examined aspects of model sensitivity found that increasing values of D led to greater spread of probability values across the probability surface and resulted in local, but not global, changes to the probability surface (Woillez et al., 2016). One potential caution about increasing D is the poor performance for both MAE and coverage probability that were observed with high values of D for large grid sizes in heterogeneous and sloping areas. However, our simulations do indicate that for small grid sizes in areas with high gradients, an increase in D should increase coverage probability with only small reduction in MAE performance.

Another option for addressing the discrepancy between probability quantiles and coverage probability is to determine which probability quantile actually corresponds to the coverage probability desired. For example, a 99% probability polygon would provide a 95% coverage probability for some treatments in the 100-m and 200-m grid sizes in the heterogeneous area. Probability values are typically so small that increasing the quantile from 99% to 99.99% would add many more grid cells to probability polygons. This option would preserve the accuracy of the reconstructed pathways. However, additional work would be needed to verify this approach, preferably in a larger study area without potentially

confounding boundary issues. In addition, this result should be confirmed for recently-developed HMM geolocation software packages such as HMMoce (Braun et al., 2017) that utilize a slightly different method for constructing the movement kernel compared to the Matlab version provided by Pedersen et al. (2008).

3.5.5 Caveats

The simulation exercises reported here are primarily aimed at providing insight into the mechanisms and relative performance of different likelihood and grid-size treatments and should not be taken as indications of specific values for precision or accuracy expected in each area. There are a number of caveats that should be kept in mind when considering this work. First, edge effects from conducting simulations in relatively small study areas may be present. However, all areas should be equally affected by edge effects, as study areas were all the same shape and size. Second, multibeam bathymetry was assumed to be known without error for simulations, but error could be present in multibeam data sets. In particular, the multibeam area in the "flat" area of Norton Sound has such little variation in depth that artifacts from the multibeam mapping may influence study area attributes such as slope and skewness. Finally, this work is a small subset of questions about model sensitivity. Other issues that should be tested include the effect of map error, different movement patterns of fish (e.g., home range or diffusive), or the effects of gaps in data.

3.6 Conclusions

The HMM geolocation model developed by Pedersen (Pedersen et al., 2008; Thygesen et al., 2009) offers a number of advantages that may make it a valuable tool for geolocation of demersal fishes in Alaskan waters and elsewhere. However, adapting the model to different applications is not always a straightforward process. It involves obtaining an understanding of habitat attributes, fish movement rates, precision and accuracy of bathymetric maps, and an accurate estimate of uncertainty in tag measurement. We anticipate that the sensitivity analysis procedures performed here will serve as a guideline for researchers planning studies of demersal fish movement in other regions based on the degree of depth heterogeneity present in their study area. We considered depth as the primary parameter in this analysis,

but many of the same considerations and conclusions should hold true for other parameters that might be used to model animal movements, such as SST, ocean heat content, or temperature-depth profiles (Braun et al., 2017).

The work we have done here is relevant to the geolocation efforts for other species as well because it illustrates the effect that different decisions about fixed model parameters have on model performance. This is increasingly relevant for fish movement research because of the availability of "black box" geolocation services provided by different companies (e.g., Argos Track & Lock or Wildlife Computers GPE3) which may not customize data likelihoods, grid size, or movement speed for individual applications. Without a full understanding of the parameterization of those models, it is difficult to interpret their results with respect to how accurately they might characterize true movement pathways and habitat use, and the relative trade-offs that they incorporate. It is unlikely that any single parameterization will perform equivalently across species and among diverse habitats.

3.7 Acknowledgments

We thank Martin Pedersen for his work developing the Hidden Markov Model for demersal fishes and making his Matlab code available to us and other researchers. We also thank Arnault Le Bris for providing valuable comments and suggestions that improved the manuscript. Funding for this research was provided by the Rasmuson Fisheries Research Center and the Pollock Conservation Cooperative Research Center at the University of Alaska Fairbanks and the North Pacific Research Board Graduate Student Research Award. Funding sources played no role in study design, analysis, or manuscript writing.

3.8 Literature cited

Andersen, K. H., Nielsen, A., Thygesen, U. H., Hinrichsen, H. H., and Neuenfeldt, S. 2007. Using the particle filter to geolocate Atlantic cod (*Gadus morhua*) in the Baltic Sea, with special emphasis on determining uncertainty. Canadian Journal of Fisheries and Aquatic Sciences, 64: 618-627.

Biais, G., Coupeau, Y., Séret, B., Calmettes, B., Lopez, R., Hetherington, S., and Righton, D. 2017. Return migration patterns of porbeagle shark (*Lamna nasus*) in the Northeast Atlantic: implications for stock range and structure. ICES Journal of Marine Science, 74: 1268-1276.

Börger, L., and Fryxell, J. 2012. Quantifying individual differences in dispersal using net squared displacement. *In* Dispersal and spatial evolutionary ecology, pp. 222 - 230. Ed. by J. Clobert, M. Baguette, T. Benton, and J. Bullock. Oxford University Press.

Braun, C. D., Galuardi, B., and Thorrold, S. R. 2017. HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. Methods in Ecology and Evolution, doi: 10.1111/2041-210X.12959.

Braun, C. D., Skomal, G. B., and Thorrold, S. R. 2018. Integrating archival tag data and a high-resolution oceanographic model to estimate basking shark (*Cetorhinus maximus*) movements in the western Atlantic. Frontiers in Marine Science, doi: 10.3389/fmars.2018.00025.

Carlson, P. R., Hooge, P., Cochrane, G., Stevenson, A., Dartnell, P., and Lee, K. 2002. Multibeam bathymetry and selected perspective views of main part of Glacier Bay, Alaska. U.S. Geological Survey, Menlo Park, CA. Open-File Report 02-391.

Cullen, A. C., and Frey, H. C. 1999. Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs. Springer-Verlag US. 336 pp.

Danielson, S., Curchitser, E., Hedstrom, K., Weingartner, T., and Stabeno, P. 2011. On ocean and sea ice modes of variability in the Bering Sea. J. Geophys. Res., 116: C12034.

Delignette-Muller, M. L., and Dutang, C. 2015. fitdistrplus: An R package for fitting distributions. Journal of Statistical Software, 64: 1-34.

Doherty, P. D., Baxter, J. M., Gell, F. R., Godley, B. J., Graham, R. T., Hall, G., Hall, J., et al. 2017. Long-term satellite tracking reveals variable seasonal migration strategies of basking sharks in the north-east Atlantic. Scientific Reports, 7: 42837. doi: 42810.41038/srep42837.

Fissel, B., Dalton, M., Felthoven, R., Garber-Yonts, B., Haynie, A., Kasperski, S., Lee, J., et al. 2016. Stock Assessment and Fishery Evaluation report for the groundfish fisheries of the Gulf of Alaska and Bering sea/Aleutian Islands area: economic status of the groundfish fisheries off Alaska, 2015. NOAA Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center.

Hanselman, D. H., Heifetz, J., Echave, K. B., and Dressel, S. C. 2014. Move it or lose it: movement and mortality of sablefish tagged in Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 72: 238-251.

Hunter, E., Metcalfe, J. D., Holford, B. H., and Arnold, G. P. 2004. Geolocation of free-ranging fish on the European continental shelf as determined from environmental variables II. Reconstruction of plaice ground tracks. Marine Biology, 144: 787-798.

Le Bris, A., Frechet, A., and Wroblewski, J. S. 2013. Supplementing electronic tagging with conventional tagging to redesign fishery closed areas. Fisheries Research, 148: 106-116.

Liu, C., Cowles, G. W., Zemeckis, D. R., Cadrin, S. X., and Dean, M. J. 2017. Validation of a hidden Markov model for the geolocation of Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences, 74: 1862-1877.

Loher, T., and Seitz, A. C. 2006. Seasonal migration and environmental conditions of Pacific halibut *Hippoglossus stenolepis*, elucidated from pop-up archival transmitting (PAT) tags. Marine Ecology Progress Series, 317: 259-271.

National Ocean Service Hydrographic Surveys 2004. H11393: NOS Hydrographic Survey, Southwest Prince William Sound, Alaska, 2004-10-19. NOAA.

Neat, F. C., Bendall, V., Berx, B., Wright, P. J., Cuaig, M. O., Townhill, B., Schon, P. J., et al. 2014. Movement of Atlantic cod around the British Isles: implications for finer scale stock management. Journal of Applied Ecology, 51: 1564-1574.

Neilson, J. D., Loefer, J., Prince, E. D., Royer, F., Calmettes, B., Gaspar, P., Lopez, R., et al. 2014. Seasonal distributions and migrations of northwest Atlantic swordfish: Inferences from integration of pop-up satellite archival tagging studies. PLoS ONE, doi: 10.1371/journal.pone.0112736.

Neuenfeldt, S., Hinrichsen, H.-H., Nielsen, A., and Andersen, K. H. 2007. Reconstructing migrations of individual cod (*Gadus morhua* L.) in the Baltic Sea by using electronic data storage tags. Fisheries Oceanography, 16: 526-535.

Nielsen, A., Bigelow, K. A., Musyl, M. K., and Sibert, J. R. 2006. Improving light-based geolocation by including sea surface temperature. Fisheries Oceanography, 15: 314-325.

NOAA NCEI 2004. Multibeam Bathymetry Database (MBBDB), W00260. NOAA National Centers for Environmental Information. doi:10.7289/V56T0JNC, accessed 17 May 2017.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. 2008. State-space models of individual animal movement. Trends in Ecology & Evolution, 23: 87-94.

Pedersen, M. W., Patterson, T. A., Thygesen, U. H., and Madsen, H. 2011. Estimating animal behavior and residency from movement data. Oikos, 120: 1281-1290.

Pedersen, M. W., Righton, D., Thygesen, U. H., Andersen, K. H., and Madsen, H. 2008. Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. Canadian Journal of Fisheries and Aquatic Sciences, 65: 2367-2377.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Seitz, A. C., Norcross, B. L., Wilson, D., and Nielsen, J. L. 2006. An evaluation of light-based geolocation for demersal fish in high latitudes. Fishery Bulletin, 104: 571-578.

Shimada, A., and Kimura, D. 1994. Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. Fishery Bulletin, 92: 800-816.

Strøm, J. F., Thorstad, E. B., Chafe, G., Sørbye, S. H., Righton, D., Rikardsen, A. H., and Carr, J. 2017. Ocean migration of pop-up satellite archival tagged Atlantic salmon from the Miramichi River in Canada. ICES Journal of Marine Science, 74: 1356-1370.

Teo, S. L. H., Boustany, A., Blackwell, S., Walli, A., Weng, K. C., and Block, B. A. 2004. Validation of geolocation estimates based on light level and sea surface temperature from electronic tags. Marine Ecology Progress Series, 283: 81-98.

Thygesen, U., Pedersen, M., and Madsen, H. 2009. Geolocating Fish Using Hidden Markov Models and Data Storage Tags. *In* Tagging and Tracking of Marine Animals with Electronic Devices, pp. 277-293. Ed. by J. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert. Springer Netherlands.

VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., and Storlie, C. 2014. SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises. R package version 1.1-221.

Viterbi, A. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory, 13: 260-269.

Webster, R. A., Clark, W. G., Leaman, B. M., and Forsberg, J. E. 2013. Pacific halibut on the move: a renewed understanding of adult migration from a coastwide tagging study. Canadian Journal of Fisheries and Aquatic Sciences, 70: 642-653.

Woillez, M., Fablet, R., Ngo, T.-T., Lalire, M., Lazure, P., and de Pontual, H. 2016. A HMM-based model to geolocate pelagic fish from high-resolution individual temperature and depth histories: European sea bass as a case study. Ecological Modelling, 321: 10-22.

Zimmermann, M., and Prescott, M. M. 2015. Smooth sheet bathymetry of the central Gulf of Alaska. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-287, 54 p. doi:10.7289/V5GT5K4F.

Table 3.1. Median treatment values for A) depth bias (m) in weighted mean reconstructed pathways, B) depth bias (m) in most probable track reconstructed pathways, and C) diffusion coefficient (km^2). Shaded cells indicate the median is significantly different (Wilcoxon rank sum, $p < 0.05$, two-tailed test) from zero for bias metrics (A and B) or from the known value of diffusion (C). Bias toward deeper depths (positive values) and diffusion medians greater than the known empirical value for simulated paths (0.012 km^2) are shaded in dark gray and bias toward shallow depths and median diffusion values less than 0.012 km^2 are shaded in light gray. Treatments are arranged by study area (heterogeneous on the left, sloping in the middle, flat on the right) and then by depth uncertainty level (low, true, or high) and then by variance specification method (W = within-cell, S = slope, and R = roughness).

	Heterogeneous									Sloping									Flat					
	Low			True			High			Low			True			High			Low		True		High	
	W	S	R	W	S	R	W	S	R	W	S	R	W	S	R	W	S	R	W	R	W	R	W	R
A) Bias: Weighted mean																								
100	-0.12	-0.13	-0.23	-0.15	-0.13	-0.27	-0.78	-0.48	-0.53	-0.04	-0.04	-0.04	-0.06	-0.06	0.00	-0.16	-0.16	-0.17	-0.01	-0.02	0.04	0.04	0.04	0.04
200	-0.10	-0.19	0.12	-0.16	0.02	0.15	-0.45	-0.53	0.26	-0.02	-0.02	-0.04	0.00	-0.04	-0.07	-0.10	-0.19	-0.13	-0.01	0.01	0.03	0.03	0.04	0.04
400	-0.35	0.31	1.85	-0.27	0.38	1.89	-0.37	0.47	2.20	-0.01	-0.02	-0.04	-0.05	-0.03	-0.04	-0.13	-0.15	-0.05	0.00	0.01	0.02	0.04	0.30	0.03
1000	0.85	0.87	0.62	1.16	0.79	0.88	0.84	0.68	0.54	-0.01	-0.16	0.02	-0.13	-0.25	0.02	-0.18	-0.22	-0.04	0.04	0.02	0.03	0.03	0.02	0.01
2000	2.84	1.86	0.57	2.84	1.90	0.57	2.74	1.64	0.57	-0.11	-0.17	-0.23	-0.09	-0.19	-0.23	-0.14	-0.32	-0.25	-0.01	0.01	0.02	0.02	0.00	0.02
B) Bias: MPT																								
100	0.06	0.07	0.38	0.02	0.04	0.26	-0.14	-0.15	0.09	0.01	0.01	-0.01	0.01	0.00	0.01	-0.02	0.01	0.01	0.00	0.01	0.00	0.01	0.05	0.03
200	0.21	0.11	0.75	0.19	0.05	0.63	0.02	0.05	0.32	0.00	-0.01	0.01	0.02	-0.01	-0.03	-0.07	-0.05	0.00	0.00	0.01	0.00	0.00	0.02	0.02
400	0.49	0.28	0.97	0.35	0.21	0.94	0.24	0.08	0.99	0.00	0.01	0.03	0.00	0.01	0.02	-0.01	-0.04	0.01	0.00	0.00	0.01	0.00	-0.01	0.00
1000	0.88	0.54	0.46	0.68	0.53	0.48	0.69	0.62	0.58	-0.02	-0.01	0.19	-0.03	0.01	0.18	-0.01	0.07	0.23	0.03	0.01	0.01	0.02	0.01	0.02
2000	1.9	1.43	0.23	1.93	1.39	0.23	2.00	1.50	0.23	0.01	-0.12	-0.03	0.01	-0.09	-0.07	-0.15	-0.17	0.08	0.00	0.00	0.02	0.00	-0.01	0.02
C) Diffusion																								
100	0.018	0.018	0.015	0.016	0.016	0.014	0.015	0.015	0.014	0.018	0.017	0.016	0.014	0.013	0.013	0.011	0.011	0.011	0.040	0.040	0.013	0.012	0.007	0.007
200	0.017	0.017	0.014	0.016	0.016	0.014	0.015	0.015	0.013	0.017	0.017	0.016	0.014	0.014	0.014	0.011	0.012	0.013	0.040	0.040	0.013	0.011	0.006	0.006
400	0.019	0.019	0.016	0.018	0.019	0.016	0.017	0.018	0.016	0.020	0.019	0.017	0.016	0.016	0.016	0.013	0.013	0.016	0.040	0.022	0.013	0.011	0.007	0.007
1000	0.033	0.036	0.028	0.033	0.035	0.028	0.030	0.034	0.027	0.036	0.032	0.026	0.031	0.029	0.025	0.024	0.024	0.022	0.039	0.020	0.019	0.017	0.008	0.009
2000	0.040	0.040	0.029	0.040	0.040	0.029	0.040	0.040	0.028	0.040	0.040	0.035	0.040	0.040	0.033	0.040	0.040	0.034	0.040	0.034	0.029	0.028	0.011	0.017

Table 3.2. Additional treatments to improve coverage probability for the 1000-m grid size (within-cell variance) in the heterogeneous study area. Treatments consist of distribution type for bathymetry grid cells, amount of extra variance added to grid cells, distribution type for depth measurement uncertainty, value of depth measurement uncertainty, and whether diffusion (D) was estimated ("Est") or fixed (value, km^2). Median treatment values for 80% coverage probability and Mean Absolute Error (MAE; weighted mean method) metrics. For each metric, difference from the median value for the within-cell variance, true uncertainty 1000 m grid size in the heterogeneous study area. Most improved treatments for each metric are highlighted in gray. Boxplots are provided in Figure S3.2-2).

TrtNo	Grid cell distribution	Grid cell extra var (m)	Depth unc distr	Depth uncertainty (m)	D (sq km)	Covg prob 80%	Diff	MAE wt mean (m)	Diff
1	Gaussian	1	Uniform	4	Est	51	-1.0	893	-6.0
2	Gaussian	2	Uniform	4	Est	51.75	-0.3	884.5	-14.5
3	Gaussian	5	Uniform	4	Est	52	0.0	868.5	-30.5
4	Lognormal	0	Uniform	4	Est	49.25	-2.8	937	38.0
5	T (1 df)	0	Uniform	4	Est	44.75	-7.3	1011.5	112.5
6	T (2 df)	0	Uniform	4	Est	45.5	-6.5	972.5	73.5
7	T (2 df)	5	Uniform	4	Est	47.5	-4.5	948	49.0
8	T (1 df)	5	Uniform	4	Est	44.5	-7.5	989.5	90.5
9	Gaussian	0	Uniform	4	0.017	52.75	0.8	823	-76.0
10	Gaussian	0	Uniform	4	0.012	51.25	-0.8	761	-138.0
11	Gaussian	0	Gaussian	1.5	Est	54.75	2.8	824.5	-74.5
12	Gaussian	5	Gaussian	1.5	Est	53.75	1.8	820.5	-78.5
13	Gaussian	0	Gaussian	1.5	Est	48.25	-3.8	921.5	22.5
14	Gaussian	0	Gaussian	1.5	Est	50.5	-1.5	903	4.0
15	Gaussian	5	Gaussian	1.5	Est	50.75	-1.3	895	-4.0
16	Gaussian	5	Gaussian	1.5	Est	48	-4.0	905.5	6.5
17	Gaussian	0	Gaussian	1.5	Est	53	1.0	856.5	-42.5
18	Gaussian	5	Gaussian	1.5	0.023	55.25	3.3	796	-103.0
19	Gaussian	5	Gaussian	3	Est	56	4.0	814.5	-84.5
20	Gaussian	0	Gaussian	3	Est	56.75	4.8	796.5	-102.5
21	Gaussian	5	Gaussian	3	0.012	54.25	2.3	720.5	-178.5
22	Gaussian	5	Gaussian	5	0.012	56.5	4.5	697.5	-201.5
23	Gaussian	0	Gaussian	3	0.017	55.75	3.8	768.5	-130.5
24	Gaussian	0	Gaussian	3	0.023	56.5	4.5	768.5	-130.5
25	Gaussian	0	Gaussian	5	0.023	56.5	4.5	804	-95.0
26	Gaussian	5	Gaussian	5	0.023	58	6.0	756.5	-142.5
27	Gaussian	5	Gaussian	5	0.040	60	8.0	809.5	-89.5

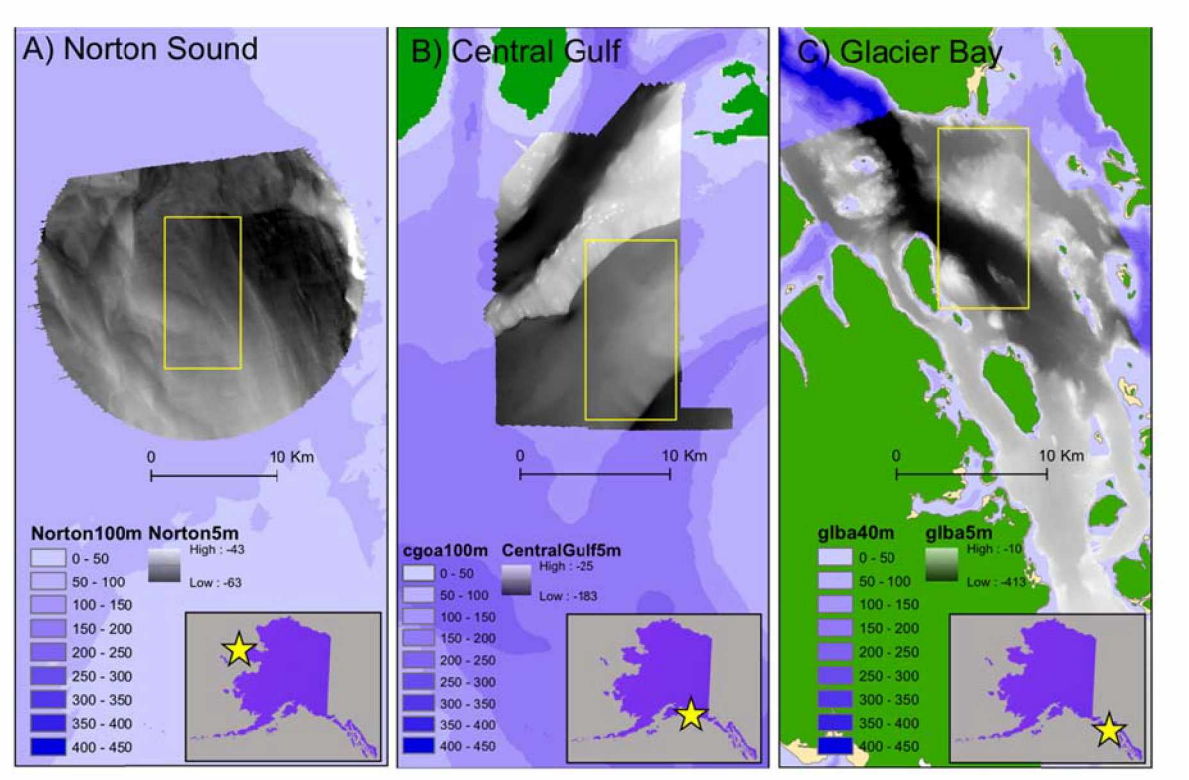


Figure 3.1. Three study areas in Alaska, USA, with different degrees of depth heterogeneity: Norton Sound (A), in northwest Alaska, has "flat" topography, the Central Gulf of Alaska (B) has "sloping" topography, and Glacier Bay (C) has "heterogeneous" topography. Areas for which multibeam sonar bathymetry data (5 m resolution) are available are shown in gray scale while larger-resolution areas adjacent to multibeam areas are shown in color. Movement paths (n=100) were simulated within yellow rectangles in each area.

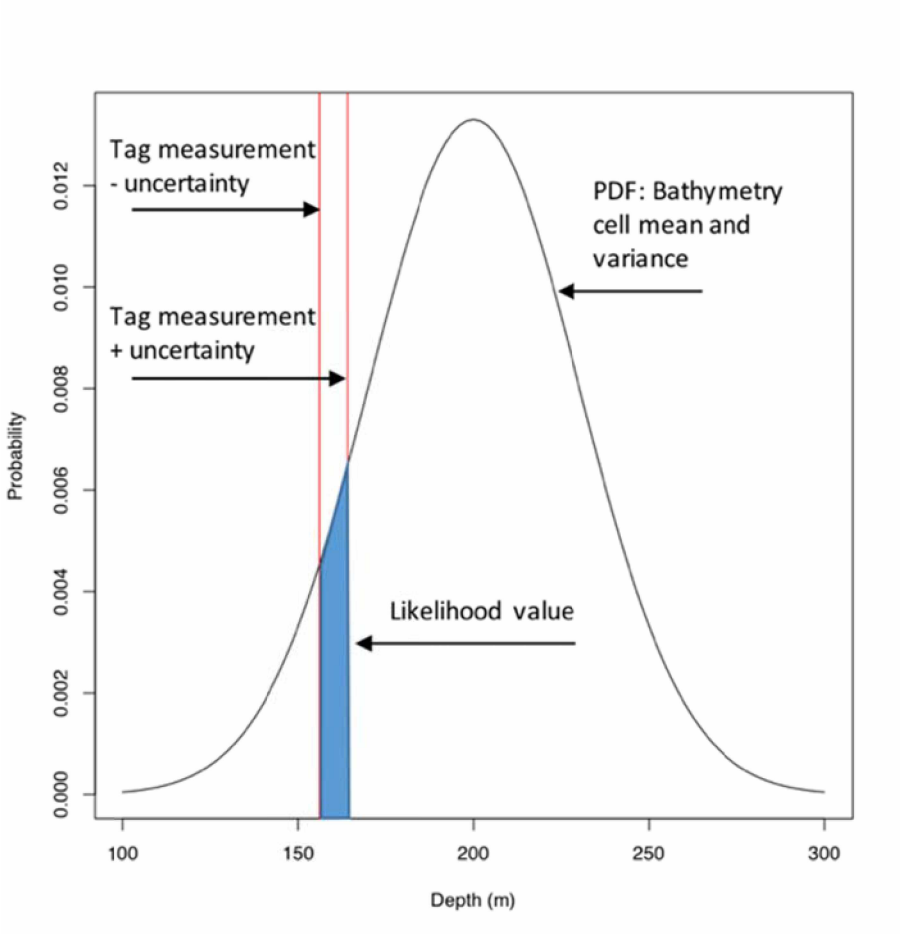


Figure 3.2. The maximum depth likelihood value (blue polygon) for each grid cell is obtained by the integration of a normal distribution derived from bathymetry mean and variance (black line) between the upper and lower limits of tag measurement uncertainty at a given time step (red lines). Study treatments consisted of three methods for determining grid-cell depth-variance and three different values of tag-measurement uncertainty.

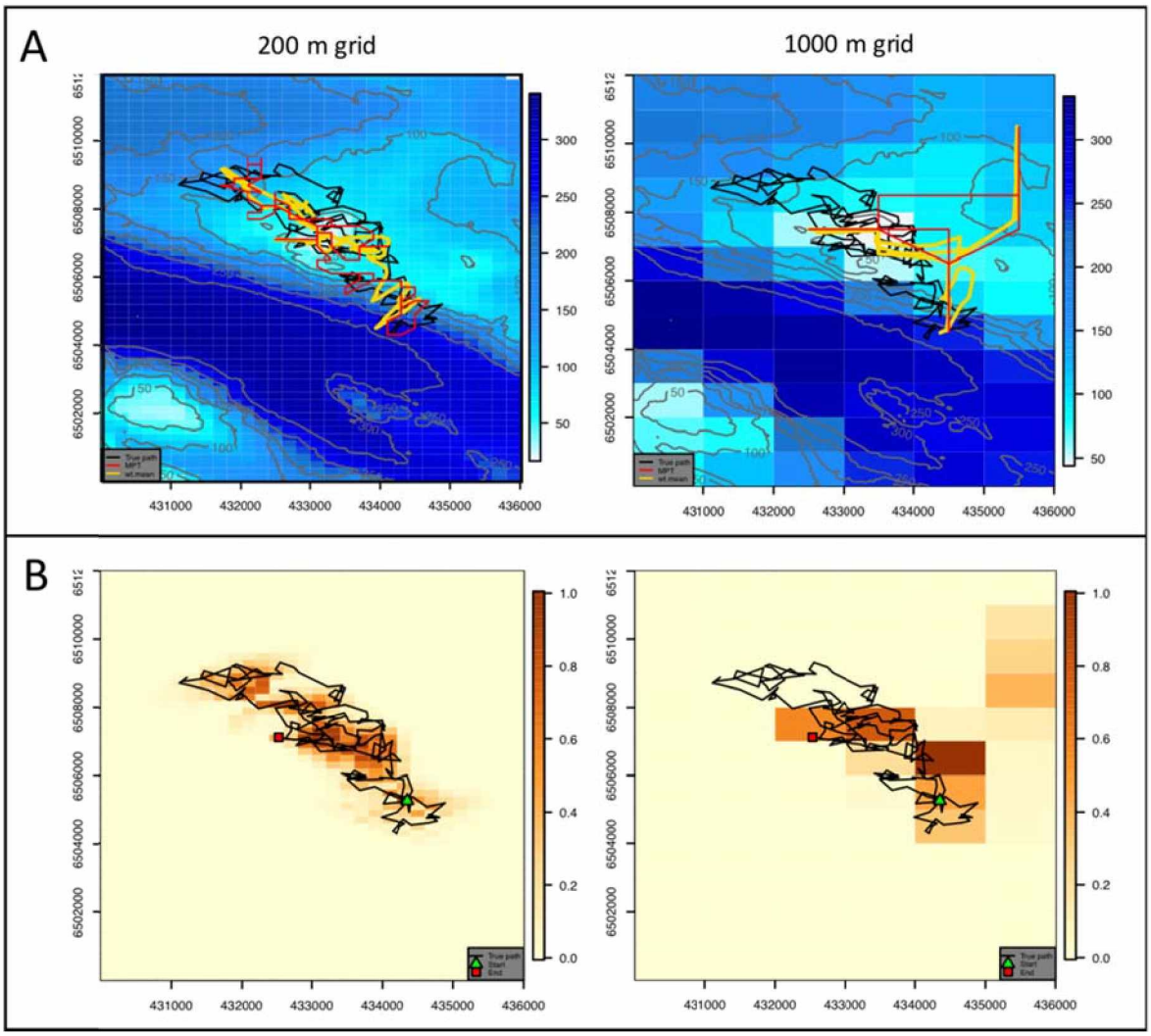


Figure 3.3. Examples of HMM results for a simulated trajectory of demersal fish movement in an area with heterogeneous bathymetry at 200-m (left) and 1000-m (right) grid sizes. A) Simulated movement paths (black lines) superimposed on study area bathymetry (gridded surface) compared to weighted mean (yellow lines) and most probable track (red lines) reconstructed pathways. B) Residency distribution (gridded surface; quantile distribution of probability in each grid cell summed over all time steps) with simulated trajectory (black lines), start locations (green triangles) and end locations (red squares). The same likelihood treatment (within-cell variance and true tag-measurement uncertainty) was used in each example.

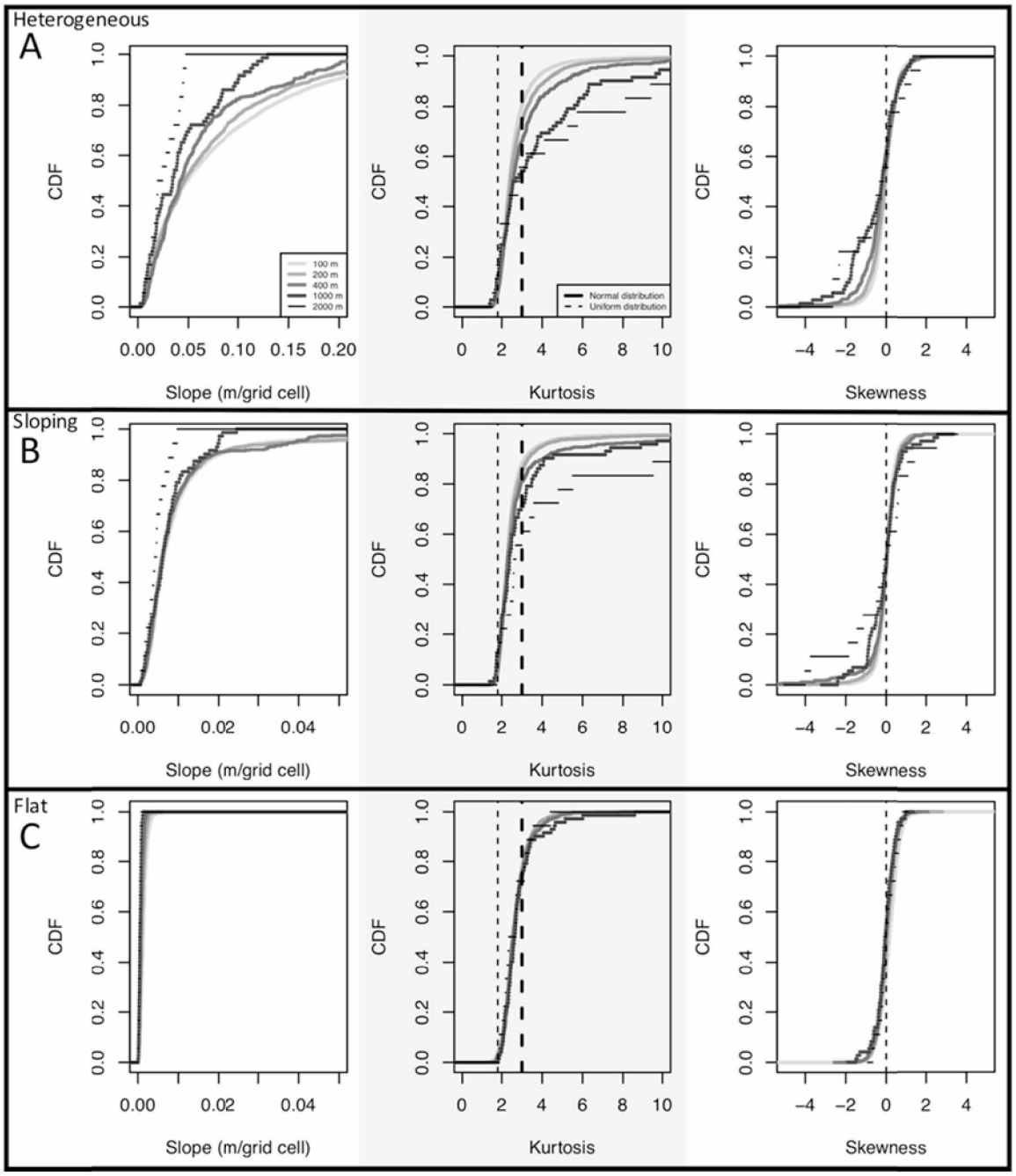


Figure 3.4. Cumulative distribution of slope (left), kurtosis (middle), and skewness (right) values for five different grid sizes in the heterogeneous (A), sloping (B), and flat (C) study areas. Kurtosis values are compared to values of 1.8 (typical of a uniform distribution, dotted line) and 3 (typical of a normal distribution, dashed line). Slope units are depth change (m) per grid cell. Note increased scale on X axis for slope in panel A compared to panels B and C.

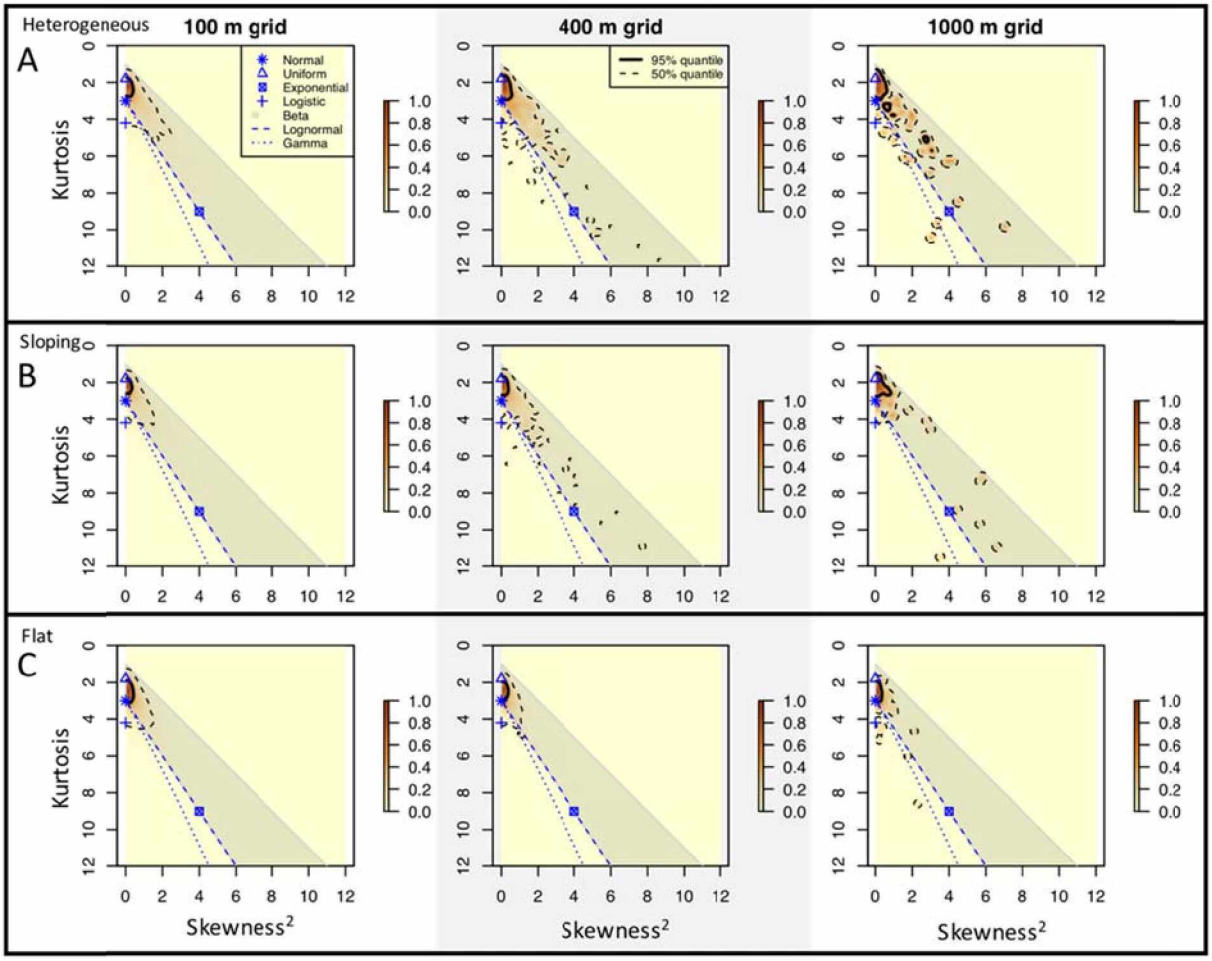


Figure 3.5. Cullen and Frey plots allow comparison of grid cell kurtosis and skewness values to known distributions for 100-m (left), 400-m (center), and 1000-m (right) grid-cell sizes from A) heterogeneous, B) sloping, and C) flat study areas. Each surface represents the quantile values of point density for grid cell kurtosis versus skewness^2 with 95% (thick black line) and 50% (thin dashed line) contours. Characteristic values of normal, uniform, exponential, and logistic distribution values are represented by blue symbols, beta distribution by a gray polygon, lognormal by a blue dashed line, and gamma by a blue dotted line.

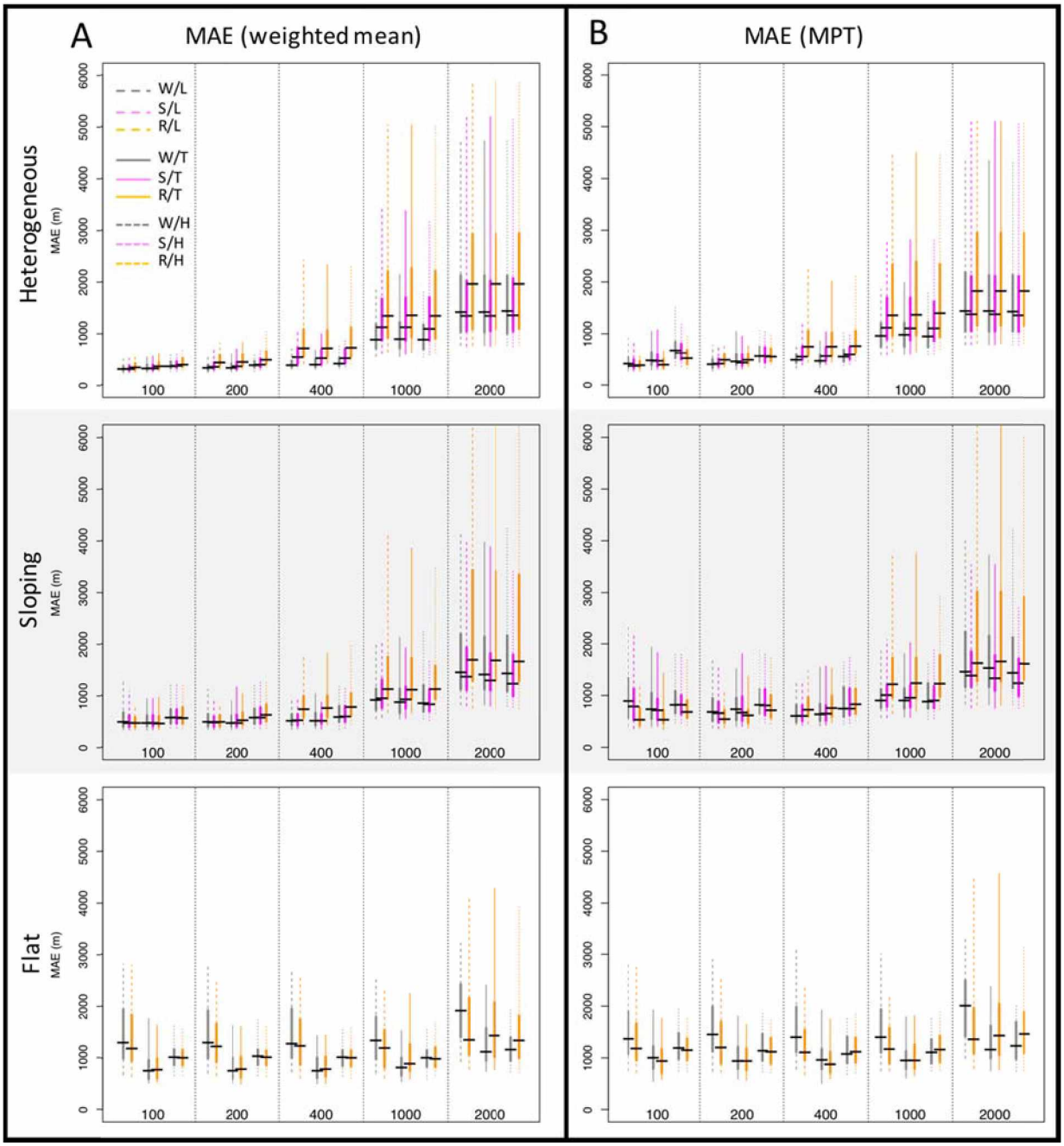


Figure 3.6. Mean Absolute Error for the weighted mean (A) and most probable track (MPT, B) methods of path reconstruction for heterogeneous (top), sloping (middle), and flat (bottom) study areas. Smaller values indicate better performance. Gray lines: within-cell variance (W), pink lines: slope variance (S), orange lines: roughness variance (R). Dashed lines: low uncertainty associated with depth measurement (L), solid lines: true depth-measurement uncertainty (T), dotted lines: high depth-measurement uncertainty (H). For each treatment, thick lines: interquartile range, thin lines: 5% - 95% quantile range, black horizontal lines: median.

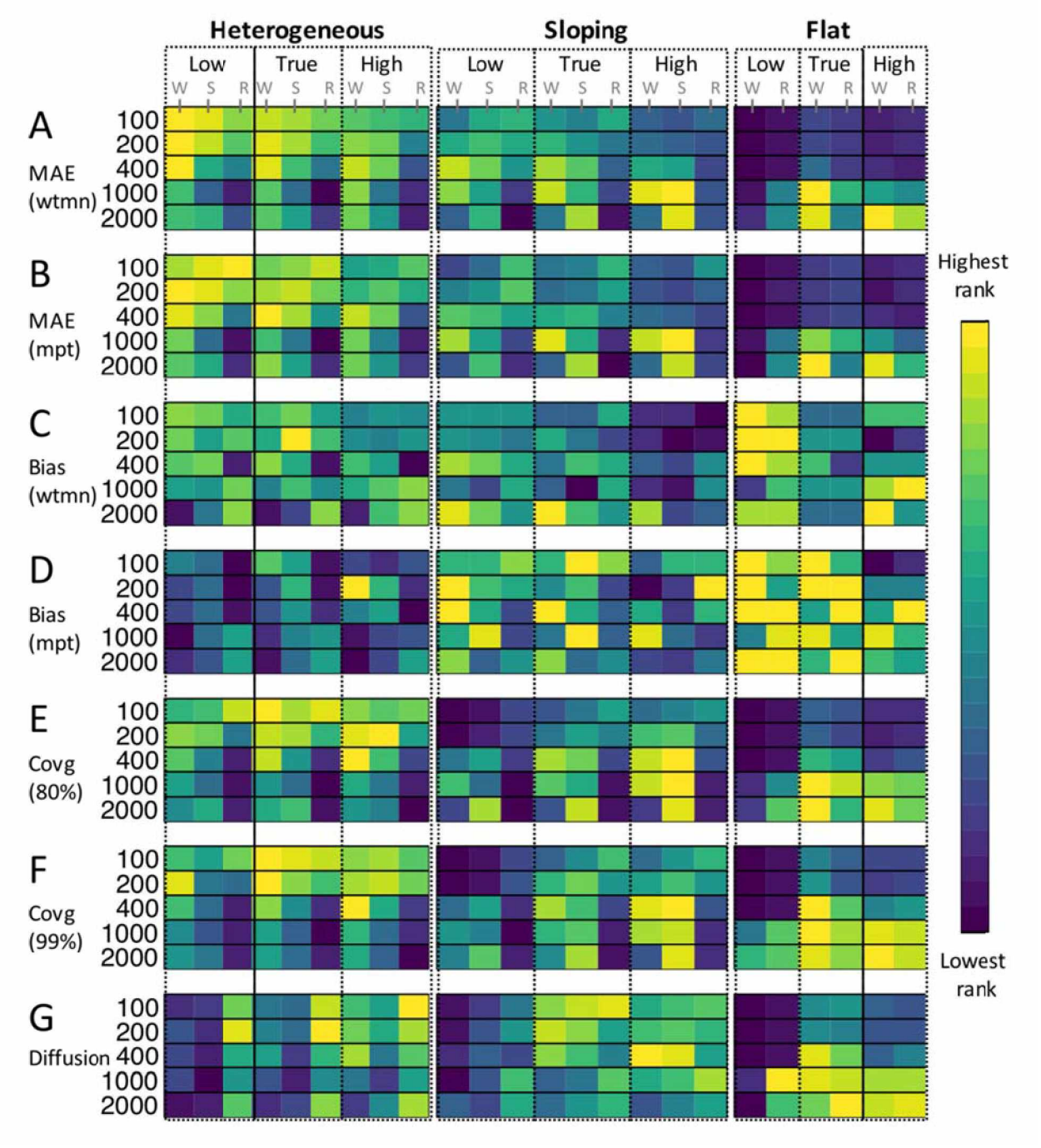


Figure 3.7. Treatment performance rankings by grid size. For each performance metric (A – G), rankings among 24 treatments (X axis) were calculated for each grid size (Y axis) and color coded (light colors have the highest ranking, i.e., the best performance). Performance metrics consisted of mean absolute error (MAE) of weighted mean (A) and most probable track (MPT, B) of reconstructed paths, depth bias for the weighted mean (C) and MPT (D) reconstructed paths, 80% (E) and 99% (F) coverage probability, and estimated diffusion coefficient (G). Treatments were grouped first by study area (heterogeneous, left; sloping, middle; flat, right), then by depth-measurement uncertainty level (low, true, or high), followed by variance-specification method (W = within-cell, S = slope, and R = roughness).

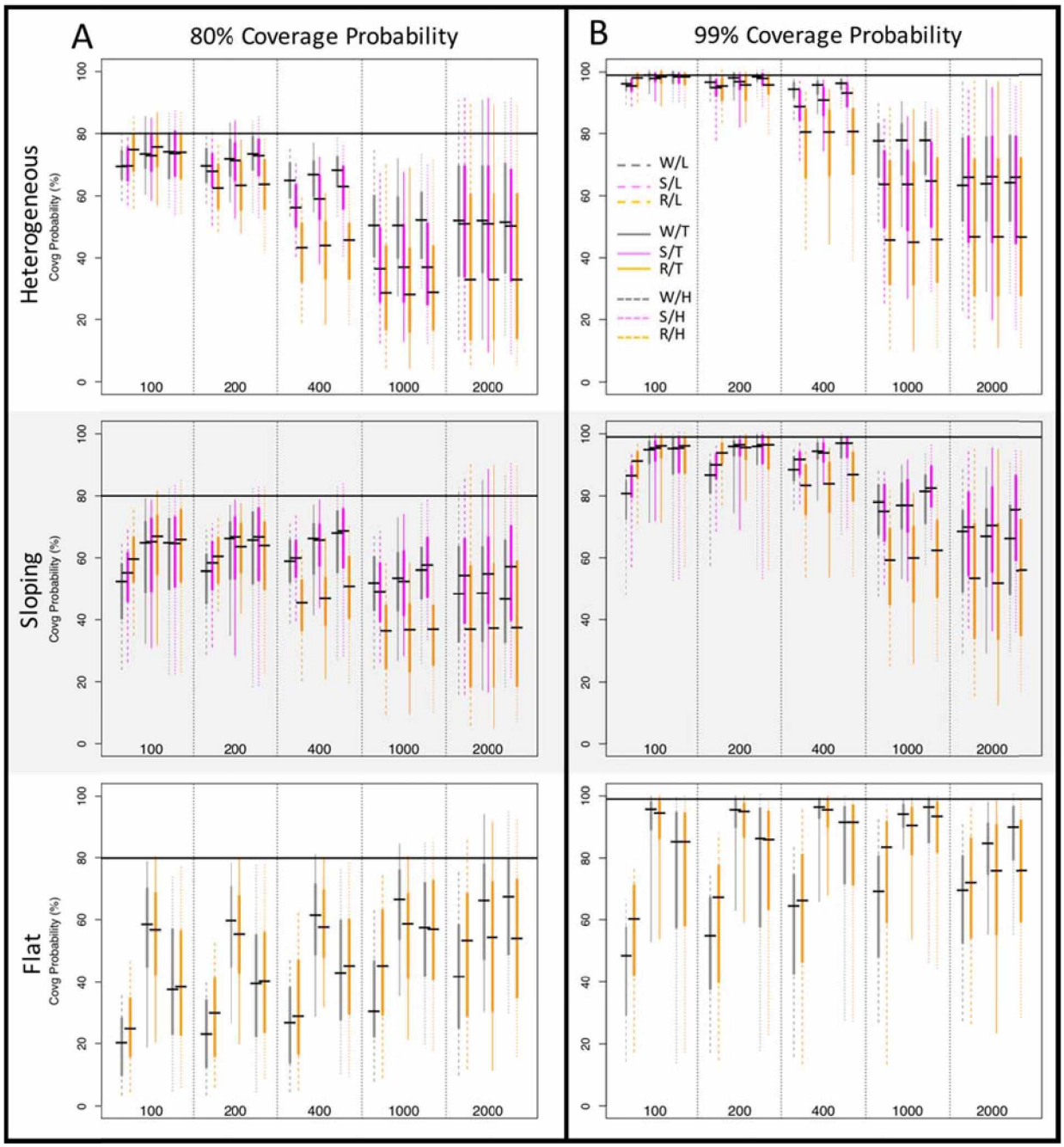


Figure 3.8. Coverage probability (proportion of known locations within estimated probability regions) for A) 80% and B) 99% probability regions in heterogeneous (top), sloping (middle), and flat (bottom) study areas. Gray lines: within-cell variance (W), pink lines: slope variance (S), orange lines: roughness variance (R). Dashed lines: low uncertainty associated with depth measurement (L), solid lines: true depth-measurement uncertainty (T), dotted lines: high depth-measurement uncertainty (H). For each treatment, thick lines: interquartile range, thin lines: 5% - 95% quantile range, black horizontal lines: median.

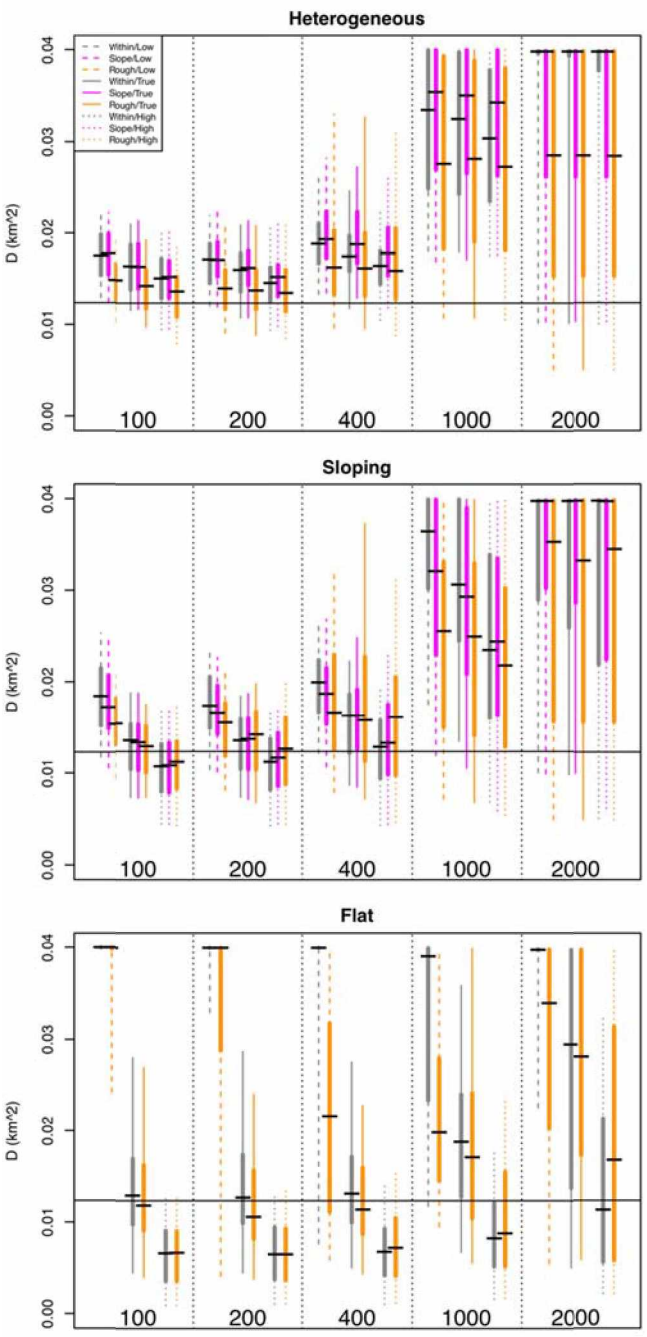


Figure 3.9. Estimated diffusion coefficient (D) in heterogeneous (top), sloping (middle), and flat (bottom) study areas. The empirical value of D calculated from simulated trajectories (0.012 km^2) is indicated by horizontal line. Gray lines: within-cell variance (W), pink lines: slope variance (S), orange lines: roughness variance (R). Dashed lines: low uncertainty associated with depth measurement (L), solid lines: true depth-measurement uncertainty (T), dotted lines: high depth-measurement uncertainty (H). For each treatment, thick lines: interquartile range, thin lines: 5% - 95% quantile range, black horizontal lines: median.

3.9 Appendices

3.9.1 Supplement S3-1. Additional information about likelihood treatments

The "slope" variance treatment was obtained by 1) determining the s.d. of depths from 5 m resolution multibeam values aggregated to form the different grid sizes used in this study (referred to as the "within-cell" variance method), 2) calculating the slope for each cell of the aggregated grids, and 3) linking the s.d. from the multibeam depths to the slope values using generalized additive model. A GAM approach was used because the relationships were not expected to be inherently linear, particularly in the area of the origin. For each grid size, s.d. was regressed against slope using a linear model, a smooth function with no basis dimension (k) specified, and smooth functions with basis dimension specifications of 3 and 5. The best model was selected by comparing AIC values and residual analysis.

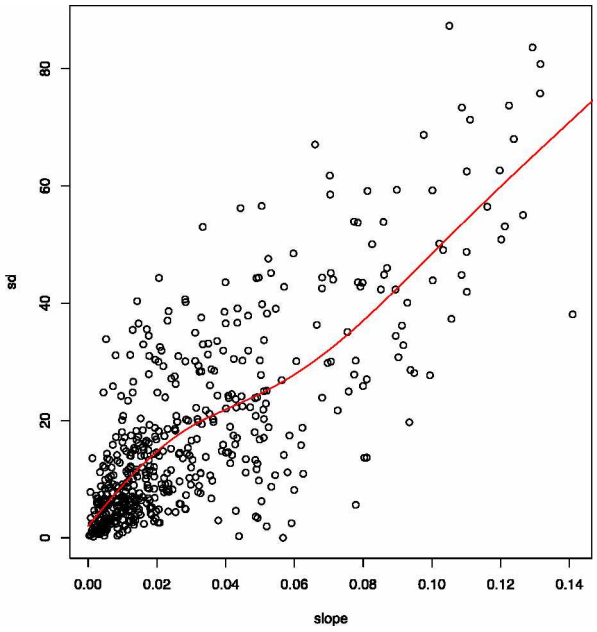


Figure S3.1-1. Example of GAM for linking s.d. to slope (the "slope" variance method) for the 1000 m grid size in the heterogeneous study area.

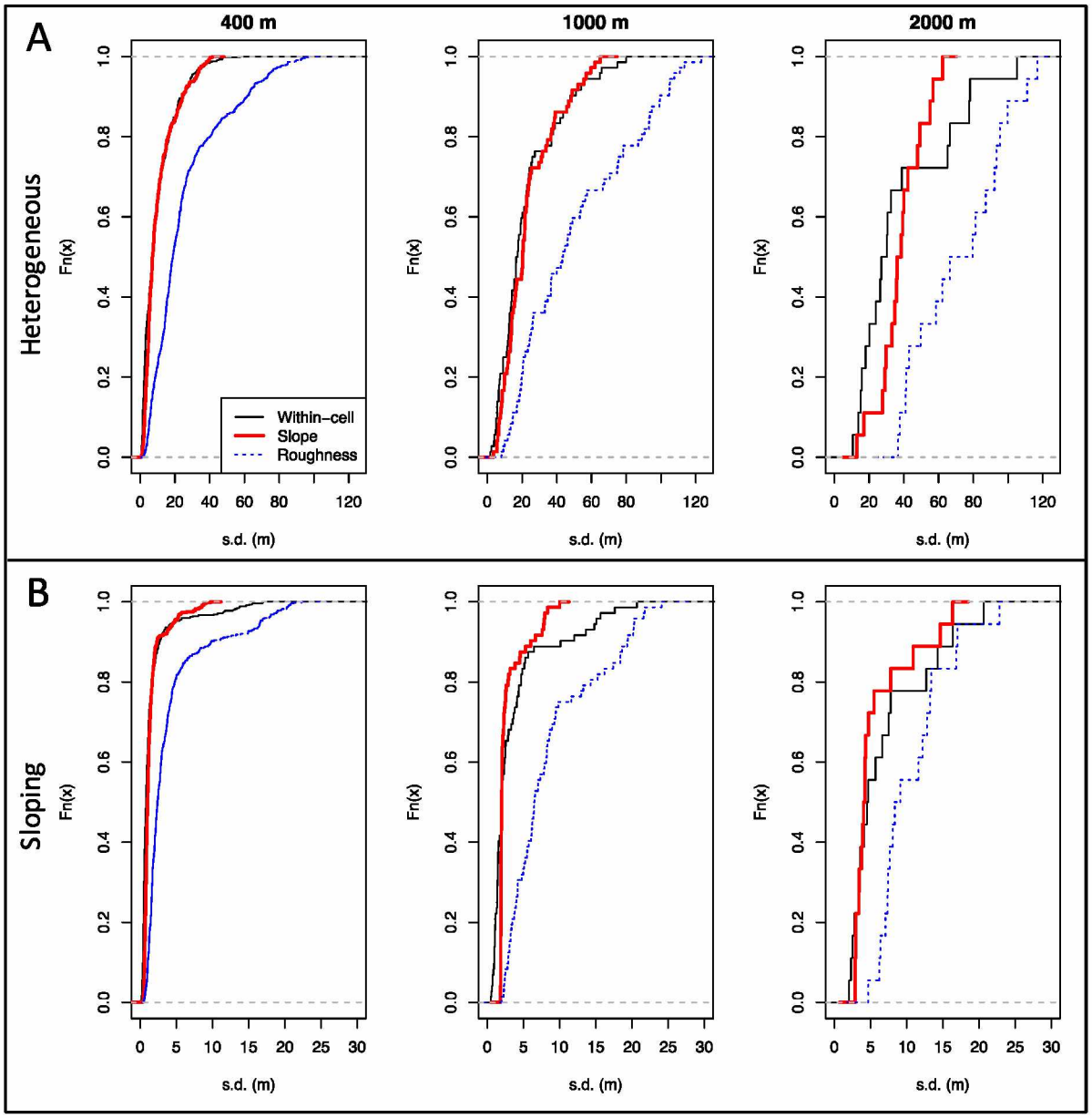


Figure S3.1-2. Cumulative Distribution Functions for "within-cell" (black lines), "slope" (thick red lines), and "roughness" (dotted blue lines) variance methods for three grid sizes in A) the heterogeneous study area, and B) the sloping study area.

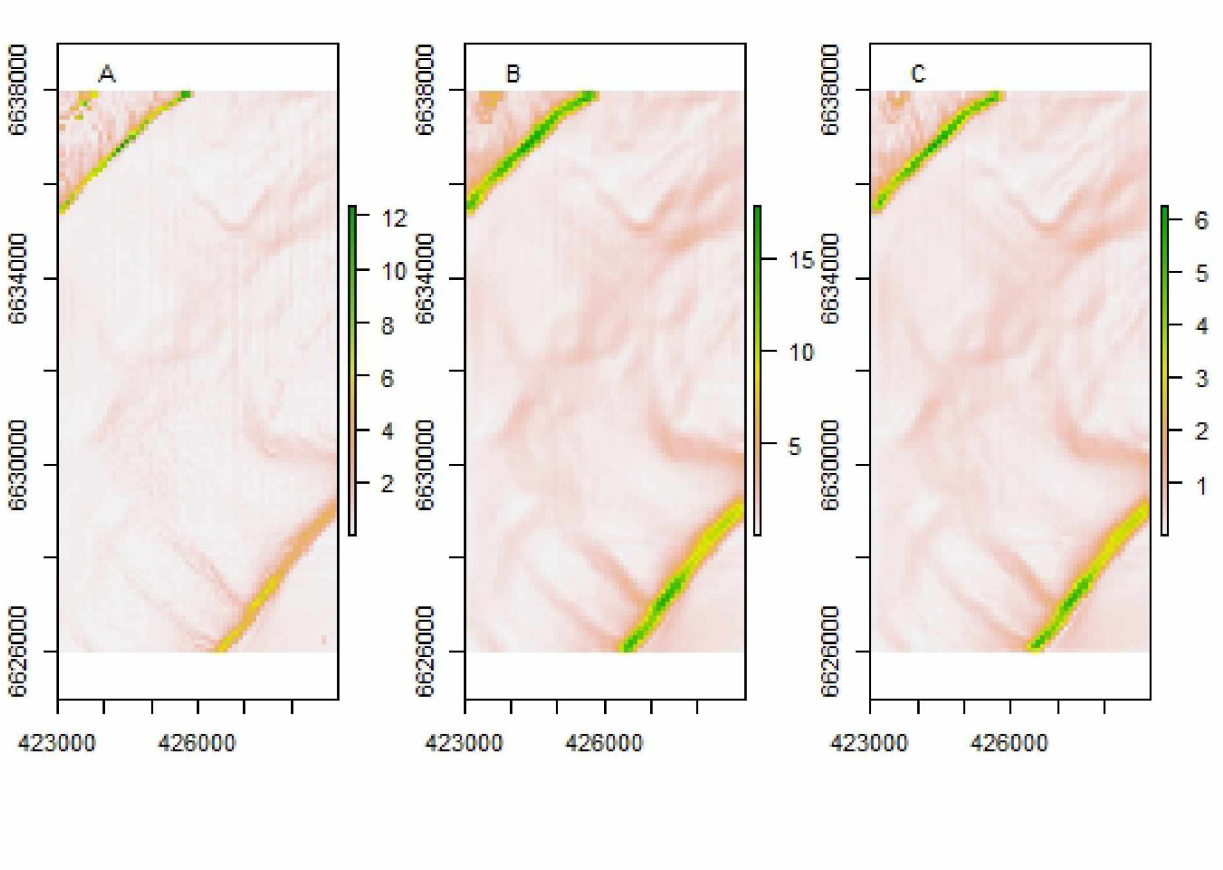


Figure S3.1-3. Three different approaches for determining variance for a 100 m grid cell in the sloping study area. A) "Within-cell", the s.d. of 5 m resolution depth measurements aggregated in each 100 m grid cell. B) "Roughness", the s.d. of the eight adjacent 100 m bathymetry cell means. C) "Slope", estimated from a Generalized Additive Model of 100 m resolution "within-cell" values (A) vs. a 100 m slope raster.

3.9.2 Supplement S3-2: Additional details about model performance

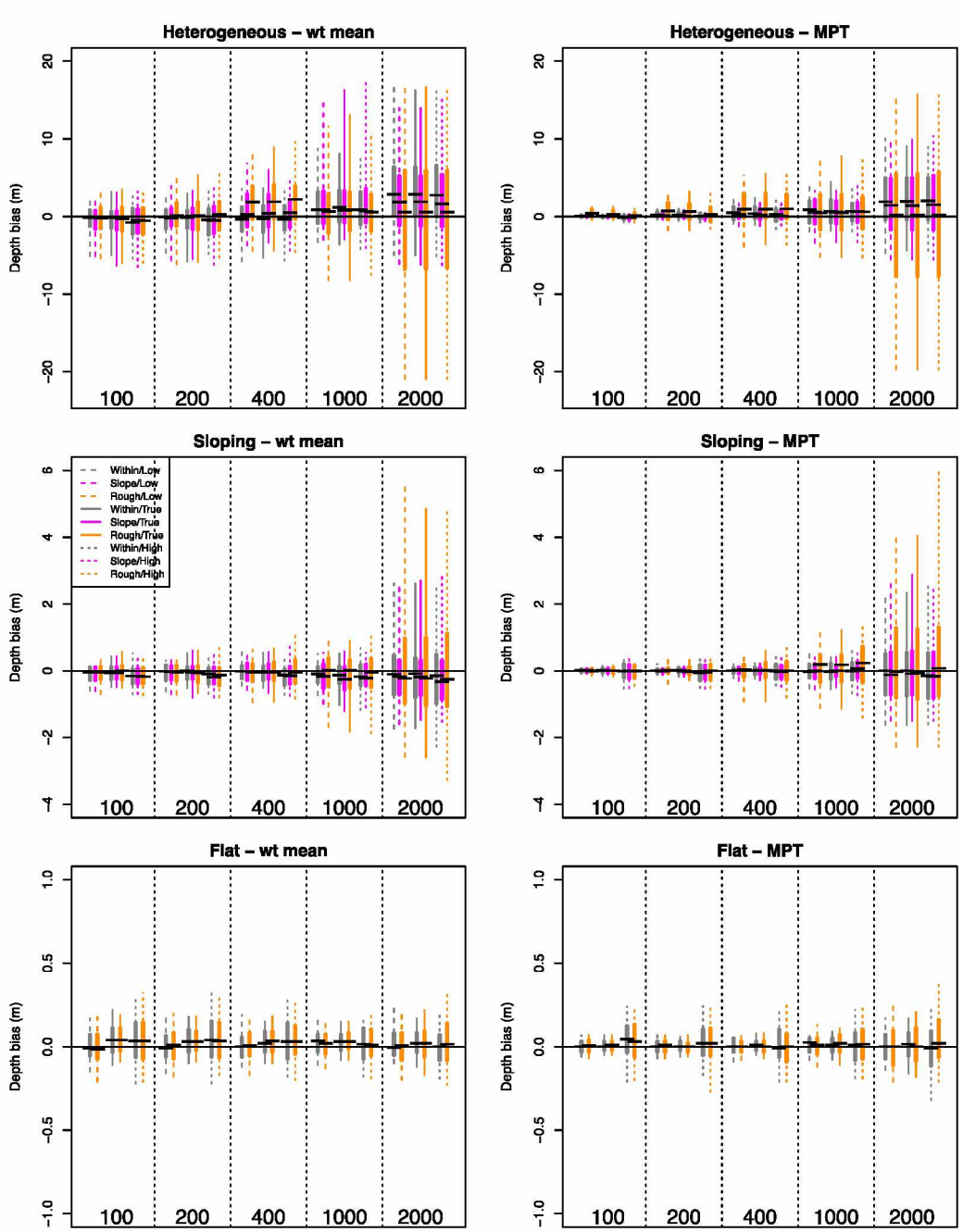


Figure S3.2-1. Bias for the weighted mean (left) and Most Probable Track (right) methods of movement path reconstruction for heterogeneous (top), sloping (middle), and flat (bottom) study areas by grid cell size. Negative values reflect a bias toward shallower depths than simulated trajectory depths. Gray lines: "within-cell" variance, pink lines: "slope" variance, orange lines: "roughness" variance. Dashed lines: "low" measurement uncertainty, solid lines: "true" measurement uncertainty, dotted lines: "high" measurement uncertainty. For each treatment, thick lines: interquartile range, thin lines: 5% - 95% quantile range.

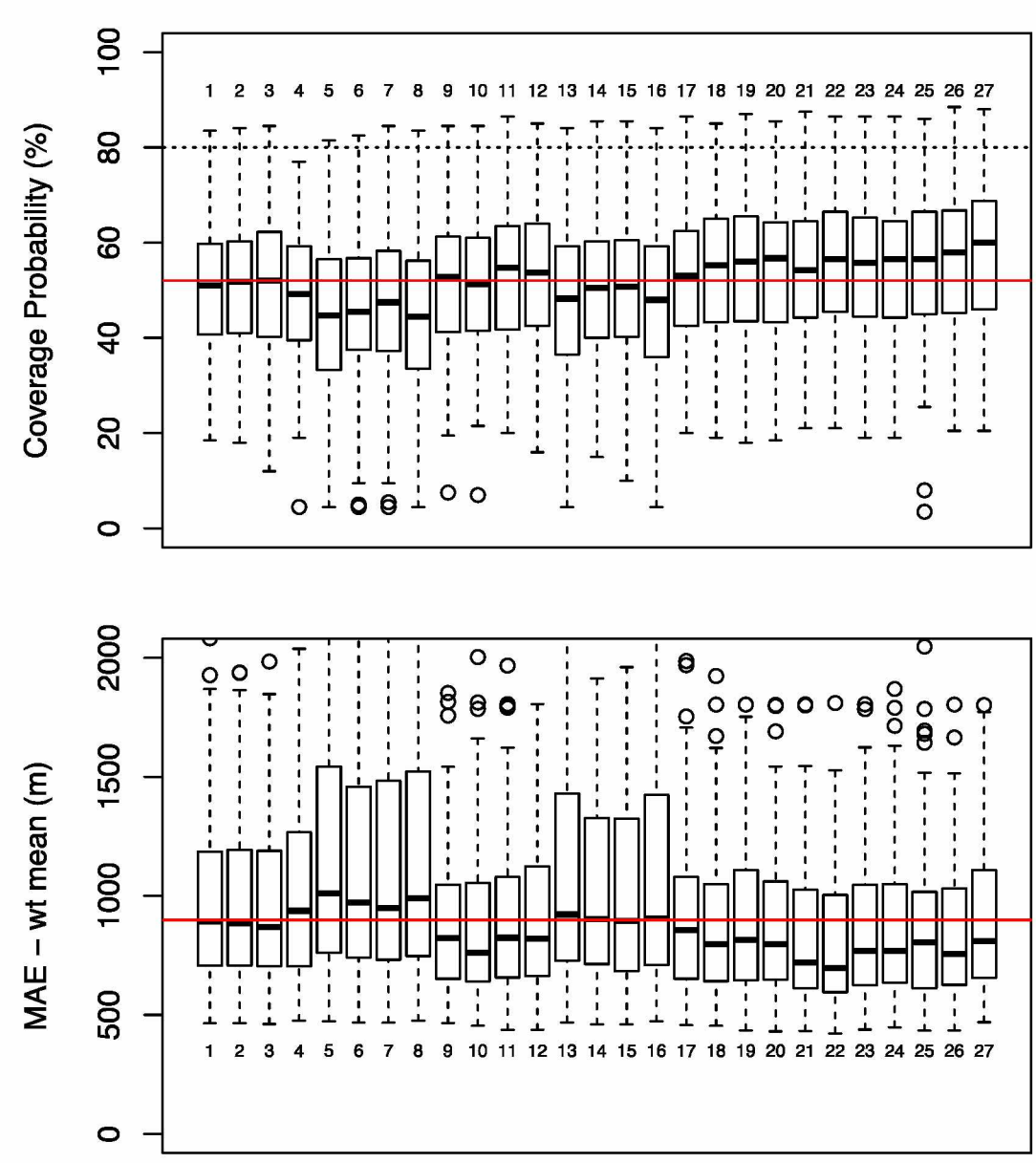


Figure S3.2-2. Mean Absolute Error and 80% coverage probability results from additional likelihood treatments for the 1000 m grid size in the heterogeneous area. Treatments are described in Table 3.2 of the main text.

Chapter 4. Potential utility of geomagnetic data for geolocation of demersal fish in the North Pacific Ocean[1]

4.1 Abstract

Archival tags that measure the Earth's magnetic field could provide a new geolocation method for demersal fishes in the North Pacific Ocean. However, the method is likely to be complicated by the presence of local magnetic field anomalies caused by geological formations, such as volcanic rock, and temporal fluctuations from solar storms at high latitudes. Therefore, the potential utility of this method in high-latitude regions that possess magnetic field anomalies, such as the North Pacific Ocean, is unknown. We assessed the theoretical utility of geomagnetic geolocation of demersal fishes in Glacier Bay National Park, USA, a high-latitude magnetic anomaly area, and hypothesized that a state-space movement model based on a combination of depth and geomagnetic data would perform better than a model based on depth alone. We developed a high-resolution (100 m) magnetic field map of the study area and assessed *in situ* tag resolution by deploying 5 geomagnetic archival tags on a stationary mooring for a period of 8 months. We simulated 1000 random walk trajectories with four levels of magnetic field measurement resolution and estimated their locations with a Hidden Markov Model using different data likelihood models. We compared results from a data likelihood model based only on depth to results from 6 data likelihood treatments that contained a combination of depth and geomagnetic data. Data likelihood treatments were composed of fine-scale or coarse-scale magnetic field maps and five possible methods of magnetic field variance specification. Geomagnetic data improved model performance for both fine-scale and coarse-scale maps when tag resolution was medium (± 500 nT) to very high (± 150 nT) and the geomagnetic variance specification was based on error between measured and mapped values instead of study area attributes such as slope or roughness. Overall, the best performance was observed for the highest resolution tag, the fine-scale map, and variance based on anomaly magnitudes. However, the coarse-scale map with a constant variance of 165 nT resulted in improvements over depth alone for all tag resolutions.

[1] Nielsen, J. K., Mueter, F., Adkison, M., Loher, T., McDermott, S., and Seitz, A. C. Formatted for submission to ICES Journal of Marine Science.

In situ testing of mooring data suggests that geomagnetic archival tags tested were capable of the equivalent of low (± 1000 nT) to medium (± 500 nT) tag measurement resolutions tested in simulations, but variation in performance was high among tags. Our results suggest that inclusion of geomagnetic data could improve geolocation of demersal fishes in the North Pacific Ocean but improvements to geomagnetic tags and additional information on magnetic field values measured at the seafloor compared to the sea surface are needed to ensure its utility.

4.2 Introduction

Demersal fish species such as Pacific halibut (*Hippoglossus stenolepis*), Pacific cod (*Gadus macrocephalus*), and sablefish (*Anoplopoma fimbria*) have extremely high economic, cultural and ecological value in the North Pacific Ocean. These fish species are capable of undertaking large-scale migrations of up to thousands of kilometers (Shimada and Kimura, 1994; Loher and Seitz, 2006; Maloney and Sigler, 2008). Detailed information on large-scale movements, such as migration timing, pathways, and the proportion of populations that migrate is important for management of these species. However, this type of information is scarce because it is difficult to obtain from conventional mark-recapture tags that provide tag release and recovery locations alone.

Electronic archival tags that collect information such as depth, temperature, and light intensity while a tagged fish is at liberty can be used to provide daily position estimates of demersal fish species over large time scales. The values recorded by the tag each day are matched to maps of the measured variables in the study area to identify the most likely location of the fish on each day, a process known as geolocation. For example, ambient light intensity measurements vary spatially and seasonally and can provide daily estimates of latitude and longitude based on the time of local noon and day length recorded by the tag each day (Musyl et al., 2001). Light-based geolocation works well for reconstructing the movement paths of highly-mobile pelagic fishes in lower latitudes, such as tunas and billfishes (Lam et al., 2008; Schaefer et al., 2011), but not high-latitude demersal fishes. Demersal fishes can occupy depths > 150 m, where light intensity is too limited for geolocation (Seitz et al., 2006) or locations where light intensity is occluded by silt or phytoplankton. The tidal method of geolocation, in which tidal amplitude

and phase from depth records of stationary fish are matched to predicted tidal amplitude and phase in the study area (Hunter et al., 2004), is often infeasible because hydrodynamic models in the North Pacific Ocean are not accurate for many locations and tidal signals in depth records can be diminished for deep water fishes. Therefore, geolocation based on maximum daily depth is a more practical method of geolocation for demersal fishes that can be assumed to be in close proximity to the seafloor at least once per day (Chapter 3). However, in areas where depth gradients are weak or for fish species that cannot be assumed to visit the seafloor on a daily basis, such as sablefish (Goetz et al., 2018), additional geolocation variables would be expected to greatly improve the accuracy and precision of geolocation estimates for these fishes in the North Pacific Ocean.

One option for an additional geolocation variable for demersal fishes in the North Pacific Ocean is provided by recently-developed electronic tags that measure the Earth's magnetic field. Because the magnetic field varies over space and can be mapped, its use for fish geolocation has been proposed (Stockhausen and Guðbjörnsson, 2009; Klimley et al., 2017). Earth's three-dimensional magnetic field can be described in terms of individual X, Y, and Z dimensions, horizontal and vertical components, or as the total magnetic field (the vector sum of all three dimensions). The magnitude of the magnetic field can be predicted at any location on earth by three-dimensional models such as the International Geomagnetic Reference Field (IGRF, Thébault et al., 2015). These global models are based on satellite measurements, vessel surveys, and magnetic observatory data. Because the main field changes slowly over time, the models need to be updated approximately every 5 years.

Two approaches for fish geolocation using magnetic field data have been introduced so far (Stockhausen and Guðbjörnsson, 2009; Klimley et al., 2017). Total magnetic field strength can serve as an indication of latitude because it increases from approximately 26,000 nanoTeslas (S.I. unit, nT) at the equator to 66,000 nT at the poles (Klimley et al., 2017). However, because magnetic field gradients and orientation vary geographically, and are not always parallel to latitude, the potential usefulness of using magnetic data to determine the latitude of tagged fish varies by geographic region. Another theoretical approach to geomagnetic geolocation features the use of separate horizontal and vertical components,

rather than the total magnetic field, because horizontal and vertical component gradients can differ within the same geographic region and thus locations may be estimated by intersecting horizontal and vertical magnetic field gradients (Stockhausen and Guðbjörnsson, 2009).

Both of these approaches, which rely on data from large-scale models of the Earth's main field, are complicated by sources of small-scale spatial and temporal variation. Small-scale spatial variation is caused by geological formations such as igneous or ferromagnetic rocks that possess their own magnetic fields. The large-scale pattern of the Earth's main field is disrupted by the magnetic field intensity of these crustal features, which can vary greatly over small spatial scales, and thus they are referred to as local magnetic anomalies. Global and regional maps of local magnetic anomalies have been produced based on vessel surveys and satellite measurements (Meyer et al., 2017). However, magnetic anomaly map resolution may be coarse compared to the spatial scale at which some local magnetic anomalies occur. Also, small-scale temporal variation results from solar and atmospheric processes. In low latitudes, charged particles in the ionosphere move more when heated by the sun during the day, resulting in diel variation in the magnetic field that occurs at a scale of 100 nT (Campbell, 2003). In high latitudes, solar storms may increase or decrease the measured magnetic field at a specific location by 2000 nT or more and can last several days (Campbell, 2003). A global network of observatories provides precise information on temporal changes at different locations; these data are freely available at www.intermagnet.org.

Because gradient strength and orientation of the main field, incidence of local magnetic anomalies, and magnitude of temporal fluctuations vary by geographic region, the potential utility of geomagnetic geolocation should be assessed regionally. In the North Pacific Ocean near Alaska, gradients in the total magnetic field run NW to SE (Figure 4.1A). For a fish moving from SW to NE, perpendicularly to the magnetic gradient orientation, gradient strength will decrease and total magnetic field values will increase. Therefore, the method would be most effective at detecting fish movement that is oriented roughly parallel to the shoreline and occurs between the Aleutian Islands and Prince William Sound, or on the eastern Bering Sea continental shelf between the eastern Aleutian Islands and the Bering

Straits. However, many local magnetic anomalies exist in Alaska, given the region's volcanic nature (Figure 4.1B). In addition, high latitudes are exposed to high temporal variability associated with the effects of solar storms (Figure 4.2). Therefore, the process of geolocation with geomagnetic data may be more difficult in Alaska (USA) compared to low-latitude regions without these phenomena.

We explored the potential value of geomagnetic geolocation for demersal fishes in the North Pacific Ocean by incorporating geomagnetic data into a discrete state-space model that can account for local magnetic anomalies. This model is based on a Hidden Markov Model (HMM) developed for the geolocation of Atlantic cod in the North Sea (Pedersen et al., 2008) and adapted for geolocation of demersal fishes in the North Pacific Ocean based primarily on depth data (Chapter 3). In this study, we expanded the HMM by developing a data likelihood model that combines depth and geomagnetic data and we hypothesized that it would perform better than a data likelihood model based on depth alone. We tested model performance under different conditions in a high-latitude magnetic anomaly area that, while small in spatial scale, allowed a mechanistic understanding of map accuracy, tag resolution, and geolocation model parameterization.

To assess the potential utility of geomagnetic geolocation with the HMM, we quantified magnetic field map accuracy, *in situ* tag resolution, and performance of different data likelihood models. We produced a fine-scale map of the magnetic field in the study area for comparison with a large-scale magnetic field map available for the broader region. We simulated movement trajectories of demersal fishes in the study area with four different values of magnetic tag measurement resolution and compared location estimation results from a model based on depth alone to results from six models that combined depth and geomagnetic data. We collected information on *in situ* tag resolution and temporal fluctuations from archival tags deployed on a stationary mooring in the study area and inferred potential utility of these tags for geolocation based on performance of simulated data with different measurement resolutions. We present the potential value of including geomagnetic data into geolocation models for demersal fishes in magnetic anomaly areas such as the North Pacific Ocean, as well as general procedures and best practices for working with geomagnetic data.

4.3 Methods

4.3.1 Study area

Our study was conducted in Glacier Bay National Park, a glacial fjord located in the northern portion of southeastern Alaska (Figure 4.1A). The heterogeneous glacial topography in the study area is composed of shallow (approximately 50 m) sills and deep (to 450 m) trenches (Figure 4.3A). Glacier Bay also has a heterogeneous geological composition formed of four distinct geological terranes produced by collision of the North American and Pacific plates (Brew et al., 1978). Magnetic anomalies in Glacier Bay are largely produced by granitic rocks from the Cretaceous and Tertiary ages (Brew et al., 1978). The magnitude of geomagnetic anomalies in Glacier Bay (range 2,000 nT) produces changes in magnetic field values over distances of a few hundred meters that are equivalent to the change in the main magnetic field over 1000 km in the vicinity of Kodiak Island, Alaska (Figure 4.1A).

4.3.2 Fine-scale magnetic field map

Because magnetic field values can vary greatly over very small spatial scales in anomaly areas, a fine-scale map of the magnetic field in the study area (Figure 4.3 B) was developed for the purposes of conducting simulations based on higher resolution magnetic field data than would be provided by larger scale maps. To obtain high-resolution magnetic field data in the study area, a GEM (Markham, Ontario, Canada) GSM-17 Overhauser magnetometer/gradiometer was attached to the bow of an aluminum vessel and data were recorded at a frequency of 1 Hz. Data were collected over the course of 7 surveys of 3 to 7 days duration conducted in the summers of 2013 and 2014. Magnetic field values recorded by stationary archival tags (SeaTagMOD, Desert Star Systems, Marina, California, USA) stationed on shore during surveys confirmed that temporal fluctuations in the study area were similar to values recorded by the nearby Sitka Magnetic Observatory (SMO, Figure 4.1A, Supplement 4-1). SMO measurements were then used to account for temporal fluctuations in the magnetic field during vessel surveys. A linear temporal trend in the main field (-85 nT/year) was removed from individual observations so that the map represented magnetic field values on July 1, 2013. All data were divided into 100 m grid cells and a mixed effects model was used to obtain cell means that accounted for autocorrelation present for each

transit through the grid cell and for each tracking trip (Supplement 4-1). Magnetic field values at locations in the study area not visited during the vessel survey were estimated by co-Kriging the vessel survey data with fine-scale aerial survey data that were available for the entire study area (Brew et al., 1978; Connard et al., 1999). Detailed methods for the construction of this map, subsequently referred to as the "fine-scale map", are available in Supplement 4-1.

4.3.3 Coarse-scale magnetic field map

A relatively coarse, large-scale map that provides information on location magnetic anomalies was obtained for the study area (Figure 4.3B). This map consisted of the North American Magnetic Anomaly Group (NAMAG, Bankey et al., 2002) anomaly grid (1-km resolution) added to main field predictions from the International Geomagnetic Reference Field (IGRF, Thébault et al., 2015). This map, referred to subsequently as the coarse-scale map, is available over a larger area in the North Pacific Ocean (Figure 4.1B) and may thus be utilized for geolocation over broader areas within the North Pacific Ocean region.

We used differences between measured and coarse-scale map values as the basis for specifying the variance of magnetic field values in each model grid cell. To determine the coarse-scale map error, we subtracted the coarse-scale map value from the value measured at each 100 m grid cell visited by the marine vessel (i.e., before the data were combined with the aerial data to create the fine-scale map of the entire study area, Figure S4.1-4A). This allowed a comparison of measured values without additional potential errors caused by co-Kriging with a different data source. We summarized map differences by coarse-scale grid cell by calculating the root mean square (rms) difference between measured and coarse-scale map values within each coarse-scale map grid cell. To determine whether map differences were larger in grid cells with higher magnetic anomalies, we tested for a relationship between rms map error vs. anomaly magnitude using a Generalized Additive Model.

4.3.4 Geomagnetic geolocation

4.3.4.1 Geolocation model

The HMM (Pedersen et al., 2008; Chapter 3) features a study area divided into discrete grid cells and ultimately estimates a probability that the tagged fish occupied a given grid cell at a given time step. Briefly, the model consists of a movement model (random walk) coupled with a data likelihood model that matches geolocation data recorded by the tag on the fish to maps of geolocation variables in the study area. At each time step, the movement model iteratively advances the probable location of the tagged fish and then the data likelihood model updates the probability that the tagged fish is located within each grid cell for that time step. Once the last geolocation record is reached, backward smoothing is conducted to re-estimate probabilities based on all geolocation records. Daily position estimates can be expressed either in terms of overall grid cell probabilities or as a single location corresponding to the mean or mode of the probability distribution across all grid cells. Uncertainty at each time step can be quantified by polygons that encompass a desired level of the probability distribution at each time step (e.g., 99%). For complete details on the HMM construction, see the dissertation appendix.

The data likelihood model for demersal fishes in the North Pacific Ocean is based primarily on the maximum daily depth recorded by a fish at each time step (Chapter 3). Because demersal fishes are assumed to be in close contact with the seafloor at least once per time step, the maximum depth can be linked to bathymetric maps of the study area. The likelihood value for each grid cell and time step is determined by integrating the probability distribution of grid cell depth values between the limits of the maximum tag depth at each time step plus and minus tag resolution (Le Bris et al., 2013). Likelihoods for additional geolocation variables, such a geomagnetic data, can be incorporated when available by cell-wise multiplication of the likelihoods for each variable at each time step. The likelihood value for geomagnetic data is calculated in the same manner as the depth likelihood (i.e., integrating the probability distribution of the total magnetic field values in each grid cell between the limits of the tag magnetic field measurement plus and minus tag measurement resolution at each time step).

4.3.4.2 Simulated trajectories

To assess the performance of different tag resolutions and HMM data likelihood models, we simulated 1000 random walks in the study area using the formula:

$$x_t = x_{t-1} + \varepsilon_{xt}\ ,\ y_t = y_{t-1} + \varepsilon_{yt}\ ,\ \text{and}\ \varepsilon_t = N(0, \sigma^2) \qquad \text{(eq 4.1)}$$

where x_t and y_t are the X and Y coordinates (here, representing changes in longitude and latitude, respectively) at time t, x_{t-1} and y_{t-1} are the X and Y coordinates at the previous time step, and ε_t is a normally distributed error with a mean of 0 and standard deviation of σ, drawn independently for the x and y components. Trajectories were simulated with 100 steps and σ = 1000 m. At each step, the depth value was extracted from the 20 m bathymetry grid and recorded; if the location fell on land, it was discarded and a new location was selected. Uniform error between +1 and -1 m was added to depth data to simulate depth measurement resolution. Magnetic field values were extracted from the 100 m fine-scale grid at each simulated location. Gaussian error for each of four standard deviation magnitudes was added to simulate four levels of tag resolution (75 nT = very high, 150 nT = high, 300 nT = medium, and 500 nT = low). To avoid effects of convoluted coastlines when using a diffusion kernel movement model (Liu et al., 2017), we altered the bathymetry for the purpose of these simulation and modeling exercises to remove inlets and convert islands to shallow areas (island and inlet cells were replaced with depth values of 5 m with random Gaussian error of s.d. = 1 m added).

4.3.4.3 Data likelihood treatments

The HMM was used to estimate locations of each of the 1000 synthesized archival data sets for 7 data likelihood model treatments. Treatments consisted of a depth-only data likelihood model and six data likelihood models that combined depth and magnetic data (Table 4.1). Data likelihood models were based on either the coarse-scale map, which has potential for large-scale application in the North Pacific Ocean, or the fine-scale map, which represents the best available magnetic field data for the study area. For the coarse-scale map, four methods of determining grid cell variance were tested (Figure 4.4). First, the "roughness" method assigns cell variance based on the standard deviation of values in all adjacent cells

and is commonly used to assign grid cell variance in other HMM applications (Le Bris et al., 2013; Braun et al., 2017; Liu et al., 2017). Second, the "slope" method assigns variance based on linking magnetic field gradients to expected variation in the grid cell from available fine-scale maps (Chapter 3). Third, the "anomaly" method assigns standard deviation values to model grid cells based on rms map error vs. the value of the NAMAG anomaly in each grid cell as described in section 4.3.3. Fourth, the "constant" method assigns the same value of standard deviation to all model grid cells; the value is derived from the 80% quantile value of rms map errors for all large-scale grid cells, as determined in section 4.3.3. For the fine-scale map, two methods of assigning variance were tested. First, the "aggregated" method consisted of the standard deviation of the high-resolution (100 m) values aggregated to form the 1 km model grid (Chapter 3); this combination of map and variance specification method represents the best available magnetic field information. Second, the "anomaly" method, described above for the coarse-scale map, was applied to the fine-scale map as well.

4.3.4.4 Model estimation

The HMM was run entirely with the R program (R Core Team, 2017) using a combination of code provided by Martin Pedersen (DTU, Denmark) and translated from Matlab to R (Chapter 3), and the R package HMMoce (Braun et al., 2017). Each of the 6 data likelihood treatments that consisted of both depth and geomagnetic data likelihood models were run with the four different tag resolution scenarios. For the movement model, we used the same value of σ that was used to create the trajectories (1000 m). The size of the diffusion (movement) kernel was 9 x 9 cells, which allowed a maximum movement of 5.66 km per time step. Pathways were reconstructed by the weighted mean method (Chapter 3) consisting of the mean location of the smoothed probability distribution at each time step.

4.3.4.5 Performance assessment

Performance of each treatment on a simulated data set was assessed by calculating the mean absolute distance between each known (simulated) location and the location estimated by the model. This quantity is referred to as the Mean Absolute Error (MAE). Box-and-whisker plots were constructed to visualize results from all data likelihood/tag resolution treatment combinations. The Wilcoxon rank sums

test was used to determine whether the median value of MAE for each depth/magnetic likelihood treatment differed from depth alone.

4.3.5 Geomagnetic archival tag resolution

To understand *in situ* archival tag magnetic measurement resolution and assess temporal change in magnetic field values due to solar storms, we deployed five geomagnetic archival tags on a stationary mooring in the study area (Figure 4.3A). Geomagnetic tags from two different manufacturers (Desert Star Systems SeaTagMOD, n=3, and Star Oddi DSTmagnetic, n=2) were rigidly attached to a mooring line at depths ranging from 134 m to 138 m from October 10, 2013 to July 1, 2014. Desert Star tags recorded measurements every 4 minutes and Star Oddi tags recorded every 20 minutes. Both types of tags recorded tri-axial magnetic field, tri-axial acceleration, depth, and temperature data. The mooring consisted of a concrete anchor and nylon mooring line and thimbles. An aluminum acoustic release Oceano 500 (iXblue, Saint-Germain en Laye, France) was mounted 2 m above the anchor and 2 m below the lower tag. Prior to deployment, a G-857 proton precession magnetometer (Geometrics, San Jose, CA, USA) was used to verify that the acoustic release did not influence the magnetic field at a distance of 2 m. Tags were spaced 1 m apart and attached to the line with plastic fasteners. Three plastic trawl floats (buoyancy 27.6 lbs each) were used for flotation and were attached 1 m above the upper tag. Tags were deployed October 9, 2013 and were recovered July 1, 2014.

To assess tag resolution and accuracy, we first calculated an offset that linked the total magnetic field value measured by the tag to the measured value of the total magnetic field at the mooring location because neither tag type recorded absolute magnetic field values. For fish geolocation, the offset is obtained by finding the difference between the daily mean of the total magnetic field values on the first (or last) day at liberty and the known value of the magnetic field at the release location (or tag recovery location). This offset is then applied to every record in the data set. To be consistent with geolocation methods for fishes, we calculated the magnetic field value at the mooring location from the fine-scale map and subtracted the mean for the first day. The offset for each tag was then added to each recorded magnetic field value.

To assess temporal trends and determine whether solar storms could be detected by the tags, we superimposed observatory data from SMO over detailed data sets. Observatory data from SMO were found to be representative of temporal changes in the magnetic field in the study area based on short-term data that were obtained from stationary tags deployed during the mapping portion of this project (Figure S4.1-2). Observatory data were obtained at a frequency of one minute and adjusted to the magnitude of the magnetic field at the mooring location from the fine-scale (100 m resolution) map by applying an offset of 206 nT. Detailed records were visually examined at weekly timescales to determine whether obvious solar storm events were evident in the magnetic field measurements recorded by the stationary tags.

Tag measurement resolution was determined by calculating the standard deviation of all total magnetic field daily means for each tag. To visualize *in situ* tag resolution and accuracy in the context of potential HMM performance and allow comparisons with HMM estimation of simulated data sets, we generated histograms of the difference between tag daily means and the fine-scale map value at the mooring location. Histograms were plotted over polygons that represented the four levels of tag measurement resolutions used to calculate likelihoods (very high = ± 150 nT, high = ± 300 nT, medium = ± 500 nT, and low = ± 1000 nT). A Wilcoxon rank sums test was performed to determine whether a bias existed in the difference between measured and mapped values for each tag.

To assess and visualize the effects of stationary tag *in situ* bias and measurement resolution, we applied the HMM to the stationary tag data as if it were obtained from a stationary fish (Liu et al., 2017). Only magnetic field data were included in the data likelihood model in order to assess the geolocation without the influence of the more-precise depth information. The best-performing data likelihood model from the simulation and the same fixed parameters (grid size, diffusion, etc.) were applied to data from all 5 stationary tags. A residence distribution that summarizes the estimated probability of location over the entire deployment period was calculated for each tag. Daily positions were calculated using the weighted mean probability of all grid cells in the study area on each day and mean absolute error (MAE) between

the mooring location and estimated daily locations was calculated. Polygons representing 99% of the probability distribution on each day were created and the area of each polygon calculated.

4.4 Results

4.4.1 Comparison of measured to coarse-scale map values

The difference between measured and coarse-scale mapped values was generally low for grid cell anomaly values between -150 and 150 nT (Figure 4.5A). There were some exceptions within this range, including several large negative values measured near the Bartlett Cove fuel dock. However, both positive and negative differences began to increase at anomaly values greater than 150, and at anomaly values greater than approximately 400 nT, differences were mostly positive (e.g., measured values tended to be greater than mapped values in grid cells with higher anomaly values).

The GAM for rms map difference vs. anomaly magnitude was significantly different from linear (smooth term: $p < 2e\text{-}16$, 4.13 estimated degrees of freedom; the basis dimension, k, was limited to 6 to prevent overfitting) and explained 34% of the deviance. The fitted model was approximately flat for anomaly values between -150 nT and 150 nT, began to increase above 150 nT, and began to level off at anomaly magnitudes of approximately 300 nT (Figure 4.5B). Based on these results, we developed three categories of variance for use in the anomaly method of variance specification by calculating the 80% quantile value of rms map differences in three categories: -150 – 150 nT, 150 – 300 nT, and > 300 nT. These quantile values were 110, 191, and 347 nT, respectively. The variance for the constant method of variance specification, 165 nT, was obtained from the 80% quantile of all rms map distances regardless of anomaly magnitude. This quantile value was chosen to represent the majority of observed rms values without being unduly influenced by higher rms values that were sometimes observed.

4.4.2 Model Performance

Increased performance compared to the depth-only likelihood model was observed with the use of both fine-scale and coarse-scale maps, but only for certain tag resolutions and variance specification methods (Figure 4.6, Table4.2). Overall, the likelihood treatments that featured the fine-scale magnetic map had the greatest performance increases compared to depth-only when tag resolution was high or very

high. The better performance of the anomaly method of variance specification compared to the aggregated method was a surprise, as the latter was expected to represent the best available map and variance information. The likelihood treatments that featured the coarse-scale magnetic map performed better than that with depth alone as long as variance specification was based on map error (e.g., either the anomaly or constant method of variance specification) and tag resolution was medium or higher. The constant method of variance specification had better performance than the anomaly method, and performance of this treatment was similar to the fine-scale map treatments for medium and high tag resolutions. For the low-resolution tags, the coarse-scale map with the constant method of variance performed better than either of the two fine-scale map treatments. However, the coarse-scale map with variance specifications based on roughness and slope had much poorer performance compared to depth alone.

Within each likelihood treatment, tag resolution had a strong effect on tag performance. Performance increased as tag magnetic field resolution increased for both fine-scale map treatments and coarse-scale map treatments with the anomaly and constant methods of variance specification. However, a trend toward better performance with increasing tag resolution was not observed for the coarse-scale map with the roughness or slope methods of variance specification.

4.4.3 Geomagnetic archival tag resolution

All five archival tags on the stationary mooring exhibited temporal patterns in detailed (4-minute interval) and daily average total magnetic field measurements that were similar to each other but not related to fluctuation in the magnetic field from solar storms (Figure 4.7). The total magnetic field measured by the SMO rarely varied by more than several hundred nT throughout the tag deployment period, though several solar storms produced variations of greater magnitudes. The range of 1-minute values measured by the SMO during the deployment period (1191 nT) was smaller than ranges observed for detailed magnetic field measurements from the stationary tags (Table 4.3). Tag DS-2 had the smallest range of detailed measurements (1496 nT) and tag SO-2 the largest (9828 nT) over the deployment period. The lowest magnetic field values recorded by SMO during the deployment period occurred during a solar

storm on February 20, 2014. This storm produced fluctuations in the magnetic field with a range of 1081 nT over a 24-hour period. Changes in total magnetic field values for some tags during this time were similar to the pattern of the solar storm (Figure 4.8; note that tag SO-2 was not included due to its much larger range of variation), and the lowest measurement recorded in the course of the entire deployment period for one of the tags (DS-1) occurred in conjunction with this storm. However, in general the daily variations in tag measurements were larger than the range of the storm and similar patterns were observed in tag data when no storms were evident in the SMO data. After a visual inspection of weekly data for all tags, we concluded that solar storm patterns could not be distinguished reliably in the stationary archival tag data.

All five stationary tags recorded daily changes in the magnetic field that appeared to be much more strongly related to tidal action than to solar storms (Figure 4.9A). Oscillations in the magnetic field measurements closely resembled the oscillations in depth due to tide (Figure 4.9B) and were greater during flood tides compared to neap. Changes in tag orientation that could result in distortion of the magnetic field (Supplement 4-2) were also associated with tidal patterns (Figure 4.9C). However, restricting magnetic field data to only those records collected when the tag was in the same orientation did not remove the oscillating magnetic field values associated with tidal variations, though it did dampen it considerably (Figure S4.2-3). A more detailed description of factors that could produce artifacts in magnetic field values measured by the tags on the stationary mooring is available in Supplement 4-2.

Geomagnetic tag measurement resolution of daily means varied within and between manufacturing type (Figure 4.7, Figure 4.10). In general, Desert Star tags were more precise than Star Oddi, with the standard deviation of daily means ranging from 159 – 303 nT (Table 4.3). Standard deviation of daily means for Star Oddi tags ranged from 543 – 2601 nT. In comparison, the standard deviation of daily means recorded by the SMO was much smaller (26.2 nT). Tags also varied in accuracy (difference from the known value at the mooring location; Table 4.3, Figure 4.10), and all tags except SO-1 exhibited significant measurement bias. For comparison with the tag resolution levels used to generate likelihoods for the geolocation simulation studies, Desert Star tag resolution would be consistent with low

(DS-1 and DS-3) to medium (DS-2) resolution (Figure 4.10). Star Oddi tag resolution ranged from low (SO-1) to so low that the data would presumably be unsuitable for geolocation (SO-2).

Four of the five stationary archival tags provided HMM-estimated locations from stationary tag magnetic field data (Figure 4.11). The data from tag SO-2 were too extreme to be found in the study area, thus the model could not function. The data would allow for estimation if a larger study area were used, however. Model reconstructed pathways for all tags appeared to wander around the study area, with the mooring location found at the edge of the estimated residency distributions of all four tags. Mean Absolute Error (MAE) of estimated locations, where lower values indicate higher accuracy, was similar for tags SO-1, DS-2, and DS-3 but much larger for tag DS-1 (Table 4.3). The mean size of daily 99% probability polygons (where lower values indicate higher precision) ranged from approximately 200 to 350 km^2.

4.5 Discussion

4.5.1 Geomagnetic geolocation with the HMM

Our simulations suggest that, despite the presence of geomagnetic anomalies in the region, geomagnetic data could improve the geolocation of demersal fishes in the North Pacific Ocean when combined with depth data in a Hidden Markov Model framework. However, the degree of potential improvement depends on the resolution and accuracy of both geomagnetic archival tags and magnetic field maps available for specific study areas as well as the variance specification method employed. Improvements in model performance relative to the use of depth data alone were observed even with coarse-scale magnetic field maps, though in general the magnitude of improvement over the depth-based model was greater for the fine-scale maps. The similar performance of the coarse-scale map with constant variance specification to the fine-scale map treatments is encouraging for the application of the coarse-scale map for geolocation over larger study areas in the North Pacific Ocean. However, it is important to note that performance could decrease compared to depth-only models if low resolution tags, poor quality maps, or mis-specified data likelihood models are employed. On the other hand, if fine-scale maps are

available for a given study area, as they are for the Glacier Bay study site, considerable improvement in geolocation over depth alone would be expected from tags with high or very high resolution.

In regions where local magnetic anomalies exist, the use of the Hidden Markov Model framework that can explicitly account for magnetic anomalies may be preferable to existing methods that rely solely on main field gradients. For example, when using a method that intersects latitude derived from the Earth's main field with longitude derived from light intensity, large geolocation errors were observed in the Galapagos region where local magnetic anomalies are prominent (Klimley et al., 2017). This method is unsuitable for demersal fishes in the North Pacific Ocean even in the absence of the errors caused by local magnetic anomalies because it requires light-based longitude data that are either too sporadic or non-existent (Goetz et al., 2018). For the method of intersecting horizontal and vertical gradients of the main field (Stockhausen and Guðbjörnsson, 2009), problems would be anticipated in anomaly areas because large-scale anomaly maps are only available for the total magnetic field, not for horizontal and vertical components, separately.

The HMM framework proposed here that explicitly incorporates magnetic anomalies has advantages beyond accounting for differences between measured and mapped values. For data likelihood models based on depth alone, increased bathymetric heterogeneity and small-scale depth gradients can improve geolocation as long as grid size is small enough to satisfy the assumption of a normal distribution in each grid cell (Chapter 3). Therefore, the increased study area heterogeneity (e.g., Figure 4.1B, areas 2 and 4) caused by magnetic anomalies could also improve geolocation performance compared to non-anomaly areas. In addition, magnetic anomalies can have large-scale patterns such as the alternating swaths of positive and negative anomalies associated with seafloor spreading (Campbell, 2003) in the Gulf of Alaska (e.g., Figure 4.1B, areas 1 and 3). Therefore, magnetic field gradients at larger spatial scales may also be stronger in anomaly areas compared to non-anomaly areas, and geolocation performance would likely increase for fish that move perpendicular to those gradients (e.g., east-west movement in area 1 or north-south movement in area 3, Figure 4.1B).

4.5.2 Geomagnetic anomaly maps

The similarities between the fine-scale and coarse-scale magnetic field maps in this study were encouraging for the use of the coarse-scale map (IGRF + NAMAG) over large areas. The key differences between map scales were 1) the magnitude of anomalies tended to be lower in the coarse scale map, and 2) man-made structures, such as the fuel dock at Bartlett Cove, were not included in the coarse-scale map. The differences in the magnitude of anomaly values can be addressed by a data likelihood model that specifies a larger variance for grid cells with larger anomaly values (e.g., Figure 4.4D). However, the presence of man-made structures is more difficult to address and may be an important source of map error. Man-made structures such as offshore petroleum platforms, wind farms, and shipwrecks can have strong magnetic field signatures due to steel structural components and electromagnetic emissions. For example, shipwrecks can produce magnetic anomalies of 10,000 nT or more (Enright et al., 2006). Both demersal and pelagic fish species can have increased abundance in the vicinity of these structures (Claisse et al., 2014; Ross et al., 2016) and behaviors such as site fidelity and homing at scales of less than 50 m to these structures have been observed (Edwards and Sulak, 2006; Lowe et al., 2009; Anthony et al., 2012; Reubens et al., 2013). Therefore, if a specific study area is known to contain major man-made structures that could attract tagged fish, efforts should be made to determine their typical magnitude and include that information on the magnetic field maps. Associations between tagged fish such as plaice, cod, and skates and man-made structures have been determined based on grid cells that represent the presence of structures such as shipwrecks and undersea cables (Wright et al., 2018), so perhaps such maps could be extended or augmented to represent potential differences between measured and mapped values that could occur in different parts of the study area.

One potential challenge that could accompany the use of the NAMAG anomaly map is the gaps in coverage that sometimes occur in certain regions such as the Bering Sea (area 5, Figure 4.1B). These gaps could be filled with information from the EMAG2v3 anomaly grid, which has world-wide coverage at a scale of 2 arc-minutes (Meyer et al., 2017). In addition, the Enhanced Magnetic Model (EMM) combines main field and anomaly data to spatial scales of approximately 50 km (NOAA NCEI, 2017).

Future studies should test model performance with these additional sources of magnetic field maps that also explicitly include anomaly information.

4.5.3 Geomagnetic variance specification methods

The differences in performance for different methods of variance specification were striking, particularly for the coarse-scale map. The data likelihood treatments that were based on the roughness and slope methods performed much worse than those based on depth alone. This result may be due to the lower values of variance produced by these methods compared to the methods that were based on map error. This is an important consideration for combining multiple types of geolocation data, and it may be advisable to set higher values of variance for geolocation variables that have lower map gradients and accuracy.

4.5.4 Geomagnetic archival tag resolution and accuracy

Although in general model performance improved with the addition of geomagnetic data, the simulation results suggest that tag resolution plays an important role in the magnitude of the improvement. The lowest resolution (± 1000 nT) did not result in performance improvement, and in one case the performance was worse than using depth alone. However, the range of anomaly values in the study area was 2000 nT, so it is possible that performance may be improved even for the lowest resolution tags in areas where gradient strength is more than twice the tag resolution. In addition, the heterogeneity of depth in the study area is greater than in other locations in the North Pacific Ocean (e.g., along shallowly-sloping expanses of continental shelf habitat), so low-resolution tags could still improve geolocation over depth alone in areas where magnetic field gradients are stronger than depth gradients.

The simulation results also suggest that very-high-resolution tags may not improve geolocation if coarse maps are used. This is an important point from the standpoint of tag manufacturing and tag expense, as high-resolution tags are more difficult to produce and would therefore cost more. In addition, a great deal of magnetic field noise (e.g., solar fluctuation) occurs below a level of approximately ±100

nT (Campbell, 2003), so producing a tag with resolution greater than this would not be expected to result in further improvement of geolocation unless the fluctuations can be taken into account.

4.5.5 *In situ* geomagnetic archival tag performance

Our results from the five geomagnetic archival tags deployed on a stationary mooring suggest several implications for geolocation performance. First, magnetic field measurement artifacts that could have been related to some aspect of attachment to the mooring line resulted in gradual increases or decreases in daily magnetic field values relative to the known value at the mooring location. Tags deployed on land that were rigidly fixed to one orientation did not exhibit temporal variation that is tidal in nature; instead, they recorded changes consistent with temporal fluctuations measured by observatories (Figure S4.1-3) or changes in temperature (Figure S4.2-4). However, given the observed changes in recorded magnetic field values resulting from changes in tag orientation (Figure S4.2-1, and discussion below), it seems likely that slight changes in tag orientation on the mooring line either with tidal action or other physical action on the mooring line over time are responsible for the tag measurements that differ markedly from observatory (SMO) data. Because the patterns in the daily means varied at much larger time-scales than tidal action (e.g., for weeks the measured value would be lower than the known value in the study area, then become higher for weeks), such temporal patterns could be mistaken for tag movement in a data set not known to be from a stationary tag. For example, extended periods of time when measured values were lower than mapped values resulted in apparent tag movement to a region of low values for tag DS-1 (Figure 4.11). However, similar periods of time when recorded values were higher than the known value at the mooring location for tags SO-1 and DS-3 did not result in as much error because positive anomaly areas were much closer to the mooring location.

The sub-daily patterns in total magnetic field data recorded by all five of the archival tags on the stationary mooring are likely related to a change in orientation or aspect of the mooring line during tide changes. The magnetic field sensors in the tags are vulnerable to a host of magnetic field distortions that can cause a change in recorded magnetic field values when the tag is rotated (Figure S4.2-1). These include hard and soft iron effects from other components in the tags, such as batteries, or errors in sensor

alignment (Emaletdinova et al., 2017; Yang et al., 2017) for which neither type of tag was calibrated. Although magnetic field sensors are also sensitive to temperature change (Figure S4.2-4), the change in temperature associated with changing tides was not great enough to explain the magnitude of daily variation in magnetic field values that was observed. The motion of ions in seawater is known to produce a magnetic field in coastal areas and could perhaps produce a tidal signal in the magnetic field data, but the magnitude of magnetic fields caused by tides is typically less than 100 nT (Campbell, 2003).

Second, the requirement for calculating an offset based on linking tag measurements to mapped values on the day of release and/or recovery can lead to tag bias that can have adverse effects on geolocation performance. In our case, the first daily mean was not necessarily representative of the true offset that could be applied with perfect knowledge of the entire data set (e.g., the mean of all daily means). This is a concern for interpretation of magnetic data from tagged fishes, as only two points in time are presumed known (tag release and tag recovery locations). Thus, the bias results reported for the tags are somewhat arbitrary, but they point to a potential systemic source of error in tag measurement that could be eliminated if the tags were to record absolute rather than relative magnitudes. In anomaly areas, an additional bias may occur due to assigning the wrong mapped value, which would then propagate the error throughout the entire data set. To ensure that this does not happen, an accurate and precise measurement of the magnetic field could be obtained at the release location.

Third, we found substantial variation in precision and accuracy among tags. This is a concern because the geomagnetic data likelihood specification relies on quantification of *in situ* tag resolution values from tags known to be stationary. To ensure that the value used will represent all possible tag resolutions for tags deployed on fish, the lowest resolution observed for stationary tags should be used to specify the likelihood. In this case, tags that have higher resolution than other tags, such as DS-2, would be penalized. The poor performance of SO-2 is troubling in this context, as it was not suitable for geolocation in the study area (values were much lower than values in the study area). An alternative approach would be to test tag precision prior to deployment so that specific resolutions could be used to specify a resolution for each tag. However, the issue of changing magnetic field values with tag

orientation should be solved before users can obtain accurate resolution values during pre-deployment testing.

4.5.6 Caveats

Although this research has provided a basis for considering the potential utility of geomagnetic geolocation in the North Pacific Ocean, several caveats should be mentioned. First, our study site was small in comparison to the scales of movement that would be expected for demersal fishes over long time periods. However, because large-scale movements are composed of a series of daily movements, geolocation models that perform well at scales of daily movement should also perform well over longer time scales. Our small study area allowed a mechanistic understanding of the characteristics of local magnetic anomaly areas and corresponding insights into ways to specify a geomagnetic data likelihood model that accounts for them, and to consider and contextualize relative tag performance within that framework, and we expect our results to be applicable over larger scales in space and time.

Second, the use of the same value of diffusion for the HMM movement model as was used to simulate the trajectories likely led to better performance than if the diffusion coefficient was estimated by the model or from the literature on fish behavior. In this study, we decided to hold the diffusion constant so that differences in performance could be attributed to data likelihood treatments or tag resolution. However, determining the sensitivity of the HMM to different values of diffusion for different applications should be investigated.

Third, our fine-scale map of the study area may contain errors and may not fully represent magnetic field values that would be measured by a tag attached to a demersal fish in the study area. The fine-scale magnetic field data used to create the map were collected as part of acoustic tracking trips for tagged fish, so the spatial and temporal distribution of survey effort to collect the data were not ideal for producing a high-resolution, comprehensive magnetic field map (Supplement 4-1). We feel the map is sufficiently accurate to represent the main features of the fine-scale anomalies for demonstrative purposes, but it is possible that the distributions of magnetic field values in the 1-km aggregated model grid cells are more skewed than suggested by our fine-scale map. Further, our magnetic data were collected at sea level,

but in anomaly areas values could be much higher on the seafloor where demersal fish are located. Thus, important work remains to compare magnetic field values at the seafloor to values at sea level as part of the validation of geomagnetic geolocation for demersal species.

Fourth, the archival tags deployed on the stationary moorings were early versions of magnetic archival tags available from each manufacturer, and current versions may perform better than the results reported in our research. The Desert Star tags were not factory-calibrated with the batteries on them, so calibration of the tag with the battery section attached was performed manually prior to deployment and it is possible that errors in calibration were introduced in this step. Therefore, our findings of *in situ* tag resolution and accuracy should not be taken as representative of current tag models. Instead, the tag resolution and accuracy results presented here should be viewed in the context of providing a basis for understanding ideal characteristics of tags and demonstrating the importance of deploying stationary tags (ideally for the entire period of time that tags are deployed on fish).

Fifth, to simplify our simulations, we focused only on spatial anomalies and ignored potential temporal fluctuations due to solar storms. Information on temporal change in tag measurement uncertainty could eventually be added in the state-space framework. For now, we recommend checking nearby magnetic observatories and simply discarding magnetic field data collected during solar storms.

4.6 Conclusions

Our simulation results from a local magnetic anomaly area suggest that geomagnetic archival tag data may improve the geolocation of demersal fishes in the North Pacific Ocean when included in a state-space geolocation model. Although our research was conducted at small spatial scales (e.g., less than 100 km) to allow mechanistic insights into model parameterization and effects of different tag magnetic field measurement resolutions, we expect that geomagnetic data will also improve geolocation at larger spatial scales (e.g., 1000 km) because 1) small-scale movement is the basis of large-scale movement, and 2) main field gradients could be detected in addition to local magnetic anomalies over larger spatial scales. These results are encouraging given the lack of information currently available on large-scale movements, particularly for deep-water fish species such as sablefish.

However, further research is needed to validate the method and more fully understand potential geolocation performance in different geographic areas. Our finding that variation in stationary tag data was far greater than magnetic field fluctuations measured by the Sitka magnetic observatory, possibly due to changes in orientation from being mounted to a flexible mooring line, suggest that calibration procedures need to be improved to ensure that tags deployed on long-term moorings provide the same level of information on *in situ* tag resolution and accuracy as tags deployed on fish. Test tags rigidly fixed to the seafloor would be expected to measure changes in magnetic field due to solar storms, but the measurement resolution recorded by fixed tags would not be expected to provide any useful information on the range of values that might be present in a moving fish until the orientation issue is solved. Information about the effect of temperature on measured magnetic field values should be quantified so that it can be incorporated into geolocation procedures if necessary. Finally, tags need to be able to record the true value of the magnetic field without relying on the calculation of an offset based on the daily average recorded on either the first or last day, which can lead to systematic bias throughout data sets. Using tags that can measure the absolute value of the magnetic field and are not affected by tag orientation or temperature change, stationary tags should be deployed throughout anomaly areas to determine the difference between magnetic field values measured on the seafloor compared to the sea surface. Such a deployment of stationary tags is necessary to quantify differences between measured and mapped magnetic field values and ensure that the variance of mapped values is specified correctly in the data likelihood model for the HMM.

4.7 Acknowledgments

We thank the Pollock Conservation Cooperative Research Center and the Rasmuson Fisheries Research Center at the University of Alaska Fairbanks for supporting graduate education and research activities of J. Nielsen. The North Pacific Research Board (Graduate Fellowship), the National Park Service, and the National Undersea Research Program provided logistical funding for this research. Dr. David Stone (University of Alaska Fairbanks) and Dr. Maurice Tivey (Woods Hole Oceanographic Institute) provided assistance with fieldwork and invaluable advice and education on geomagnetic

research methods. We thank Jim De la Bruere and crew of R/V Medeia (Alaska Department of Fish and Game) for deploying the stationary mooring. Chad Soiseth and Craig Murdoch (National Park Service, Glacier Bay National Park) provided logistical support as well as advice and feedback on the manuscript. We thank Mark Evans, Kevin Siwicke, Michael Courtney, Anne Beaudreau, Gordon Kruse, Jessica Glass, John Rodstrom, Amanda Compton, Thomas Farrugia, Maurice Tivey, David Stone, Barbara Stone, Alan Steffert, Craig Murdoch, Chad Soiseth, Julien Appignanii, and Katy Rayfield for their assistance with fieldwork. Geomagnetic data used in this study were obtained from magnetic observatories operated by the U.S. Geological Survey (USGS, geomag.usgs.gov).

4.8 Literature cited

Anthony, K. M., Love, M. S., and Lowe, C. G. 2012. Translocation, homing behavior and habitat use of ground fishes associated with oil platforms in the East Santa Barbara Channel, California. Bulletin of the Southern California Academy of Sciences, 111.

Bankey, V., Cuevas, A., Daniels, D., Finn, C. A., Hernandez, I., Hill, P., Kucks, R., et al. 2002. Digital data grids for the magnetic anomaly map of North America. Open-File Report 02-414.

Braun, C. D., Galuardi, B., and Thorrold, S. R. 2017. HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. Methods in Ecology and Evolution, doi: 10.1111/2041-210X.12959.

Brew, D. A., Johnson, B. R., Grybeck, D., Griscom, A., Barnes, D. F., Kimball, A. L., Still, J. C., et al. 1978. Mineral resources of the Glacier Bay National Monument wilderness study area, Alaska. USGS Open-file Report 78-494.

Campbell, W. H. 2003. Introduction to geomagnetic fields, Cambridge University Press, Cambridge, U.K.

Claisse, J. T., Pondella, D. J., Love, M., Zahn, L. A., Williams, C. M., Williams, J. P., and Bull, A. S. 2014. Oil platforms off California are among the most productive marine fish habitats globally. Proceedings of the National Academy of Sciences, 111: 15462.

Connard, G. G., Saltus, R. W., Hill, P. L., Carlson, L., and Milicevic, B. 1999. Alaska digital aeromagnetic database description. Open-file Report 99-0503.

Edwards, R. E., and Sulak, K. J. 2006. New paradigms for Yellowfin tuna movements and fish distributions – implications for the Gulf and Caribbean region. Proceedings of the Gulf and Caribbean Fisheries Institute, 57, pp. 283-296.

Emaletdinova, L. Y., Gainutdinova, A. V., and Gainutdinova, T. Y. 2017. An algorithm for calibrating the three-axis magnetometer. Russian Aeronautics, 60: 134-140.

Enright, J. M., Gearhart II, R., Jones, D., and Enright, J. 2006. Study to conduct National Register of Historic Places evaluations of submerged sites on the Gulf of Mexico outer continental shelf. U.S. Department of the Interior, Minerals Management Service, Gulf of Mexico OCS Region, New Orleans, LA. OCS Study MMS 2006-036. 136 pp.

Goetz, F. W., Jasonowicz, A. J., and Roberts, S. B. 2018. What goes up must come down: Diel vertical migration in the deep-water sablefish (*Anoplopoma fimbria*) revealed by pop-up satellite archival tags. Fisheries Oceanography, 27: 127-142.

Hunter, E., Metcalfe, J. D., Holford, B. H., and Arnold, G. P. 2004. Geolocation of free-ranging fish on the European continental shelf as determined from environmental variables II. Reconstruction of plaice ground tracks. Marine Biology, 144: 787-798.

Klimley, A. P., Flagg, M., Hammerschlag, N., and Hearn, A. 2017. The value of using measurements of geomagnetic field in addition to irradiance and sea surface temperature to estimate geolocations of tagged aquatic animals. Animal Biotelemetry, 5: 19.

Lam, C. H., Nielsen, A., and Sibert, J. R. 2008. Improving light and temperature based geolocation by unscented Kalman filtering. Fisheries Research, 91: 15-25.

Le Bris, A., Frechet, A., and Wroblewski, J. S. 2013. Supplementing electronic tagging with conventional tagging to redesign fishery closed areas. Fisheries Research, 148: 106-116.

Liu, C., Cowles, G. W., Zemeckis, D. R., Cadrin, S. X., and Dean, M. J. 2017. Validation of a hidden Markov model for the geolocation of Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences, 74: 1862-1877.

Loher, T., and Seitz, A. C. 2006. Seasonal migration and environmental conditions of Pacific halibut *Hippoglossus stenolepis*, elucidated from pop-up archival transmitting (PAT) tags. Marine Ecology Progress Series, 317: 259-271.

Lowe, C. G., Anthony, K. M., Jarvis, E. T., Bellquist, L. F., and Love, M. S. 2009. Site fidelity and fovement patterns of groundfish associated with offshore petroleum platforms in the Santa Barbara Channel. Marine and Coastal Fisheries, 1: 71-89.

Maloney, N. E., and Sigler, M. F. 2008. Age-specific movement patterns of sablefish (*Anoplopoma fimbria*) in Alaska. Fishery Bulletin, 106: 305-316.

Meyer, B., Chulliat, A., and Saltus, R. 2017. Derivation and error analysis of the Earth Magnetic Anomaly Grid at 2 arc min Resolution Version 3 (EMAG2v3). Geochemistry, Geophysics, Geosystems, 18: 4522-4537.

Musyl, M. K., Brill, R. W., Curran, D. S., Gunn, J. S., Hartog, J. R., Hill, R. D., Welch, D. W., et al. 2001. Ability of archival tags to provide estimates of geographical position based on light intensity. *In* Electronic tagging and tracking in marine fisheries, pp. 343–367. Ed. by J. R. Sibert, and J. L. Nielsen. Kluwer Academic Publs, Dordrecht, The Netherlands.

NOAA NCEI 2017. Enhanced Magnetic Model. NOAA National Centers for Environmental Information.

Pedersen, M. W., Righton, D., Thygesen, U. H., Andersen, K. H., and Madsen, H. 2008. Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. Canadian Journal of Fisheries and Aquatic Sciences, 65: 2367-2377.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Reubens, J. T., Pasotti, F., Degraer, S., and Vincx, M. 2013. Residency, site fidelity and habitat use of Atlantic cod (*Gadus morhua*) at an offshore wind farm using acoustic telemetry. Marine Environmental Research, 90: 128-135.

Ross, S. W., Rhode, M., Viada, S. T., and Mather, R. 2016. Fish species associated with shipwreck and natural hard-bottom habitats from the middle to outer continental shelf of the Middle Atlantic Bight near Norfolk Canyon. Fishery Bulletin, 114: 45-47.

Schaefer, K. M., Fuller, D. W., and Block, B. A. 2011. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the Pacific Ocean off Baja California, Mexico, determined from archival tag data analyses, including unscented Kalman filtering. Fisheries Research, 112: 22-37.

Seitz, A. C., Norcross, B. L., Wilson, D., and Nielsen, J. L. 2006. An evaluation of light-based geolocation for demersal fish in high latitudes. Fishery Bulletin, 104: 571-578.

Shimada, A., and Kimura, D. 1994. Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. Fishery Bulletin, 92: 800-816.

Stockhausen, H., and Guðbjörnsson, S. 2009. The earth's geomagnetic field and geolocation of fish: first results of a new approach. CM 2009/B:19.

Thébault, E., Finlay, C. C., Beggan, C. D., Alken, P., Aubert, J., Barrois, O., Bertrand, F., et al. 2015. International Geomagnetic Reference Field: the 12th generation. Earth, Planets and Space, 67: 1-19.

Wright, S.R., C.P. Lynam, D.A. Righton, J. Metcalfe, E. Hunter, A. Riley, L. Garcia, P. Posen, K. Hyder, and D. 2018. Structure in a sea of sand: fish abundance in relation to man-made structures in the North Sea. ICES Journal of Marine Science, http://dx.doi.org/110.1093/icesjms/fsy1142.

Yang, D., You, Z., Li, B., Duan, W., and Yuan, B. 2017. Complete tri-axis magnetometer calibration with a gyro auxiliary. Sensors (Basel, Switzerland), 17: 1223.

Table 4.1. Data likelihood treatments compared with simulation study. Treatment name, type of magnetic field map used (aggregated fine-scale = fine-scale map aggregated to 1 km, coarse-scale = NAMAG + IGRF), and variance specification method (see Figure 4.4).

Treatment name*	**Magnetic map**	**Magnetic variance**
Depth-only	---	---
Fine/Aggregated	Aggregated fine-scale	Aggregated
Fine/Anomaly	Aggregated fine-scale	Anomaly
Coarse/Anomaly	Coarse-scale	Anomaly
Coarse/Constant	Coarse-scale	165 nT for all cells
Coarse/Rough	Coarse-scale	Roughness
Coarse/Slope	Coarse-scale	Slope

Table 4.2. Performance as measured by mean absolute error of depth and magnetic data likelihood models compared to depth-only likelihood. Likelihood treatment, tag resolution, p-value for Wilcoxon rank sums test, and median, 95% quantile, and maximum MAE values for each treatment/tag resolution combination. Median MAE values significantly ($p < 0.05$) better (lower) than the depth-only treatment are highlighted in yellow, while treatments significantly worse (larger) than the depth-only treatment are highlighted in blue.

Treatment name	Tag res (s.d.)	Wilcoxon p-value	Median MAE (m)	95% MAE	Max MAE
Depth-only	NA	NA	1705	3767	8206
Fine/Aggregated	75	<2e-16	1240	2894	8504
	150	3.3e-16	1500	3741	8855
	300	1.5e-07	1571	3297	8025
	500	<2e-16	2160	7145	20754
Fine/Anomaly	75	<2e-16	1286	2331	7615
	150	<2e-16	1404	2747	7990
	300	1.2e-10	1544	3273	8151
	500	0.3	1673	3742	9028
Coarse/Anomaly	75	<2e-16	1478	3312	8368
	150	1.6e-14	1513	3251	8388
	300	0.0017	1624	3597	8214
	500	0.99	1685	3757	9117
Coarse/Constant	75	<2e-16	1482	3035	12372
	150	1.8e-14	1508	3218	8208
	300	2.3e-07	1574	3442	8251
	500	0.011	1634	3592	8149
Coarse/Roughness	75	<2e-16	2060	6806	29819
	150	1.6e-08	1908	5511	16800
	300	<2e-16	2782	7839	20698
	500	<2e-16	2378	6387	28454
Coarse/Slope	75	<2e-16	2139	18022	36033
	150	7.1e-14	1984	14190	36728
	300	<2e-16	3069	16558	38423
	500	<2e-16	2457	9392	39173

Table 4.3. *In situ* geomagnetic archival tag resolution. Tag mooring depth, offset that links tag measured total magnetic field values to mapped value at mooring location, range of detailed (4-minute resolution) total magnetic field measurements, standard deviation of total magnetic field daily means, median difference between daily means and mapped value, Wilcoxon rank sums p-value for detecting bias, mean absolute error (MAE) between mooring and estimated locations from Hidden Markov Model (HMM), and mean size of daily error estimates from HMM. No HMM results were available for tag SO-2 because values measured were outside of the range of values in the study area.

Tag ID	**Depth (m)**	**Magnetic offset (nT)**	**Range detailed (nT)**	**S.D. daily means (nT)**	**Median diff (nT)**	**Wilcoxon p-value**	**HMM MAE (km)**	**HMM Mean error poly size (km^2)**
DS-1	134	-4051	2186	303	-178	1.95e-14	14.3	242
SO-1	135	-4808	1939	543	-38	0.4879	5.3	278
DS-2	136	365	1496	159	-271	< 2.2e-16	6.3	195
DS-3	137	-2394	2020	263	183	< 2.2e-16	4.9	349
SO-2	138	-4405	9828	2602.1	-4126	< 2.2e-16	NA	NA

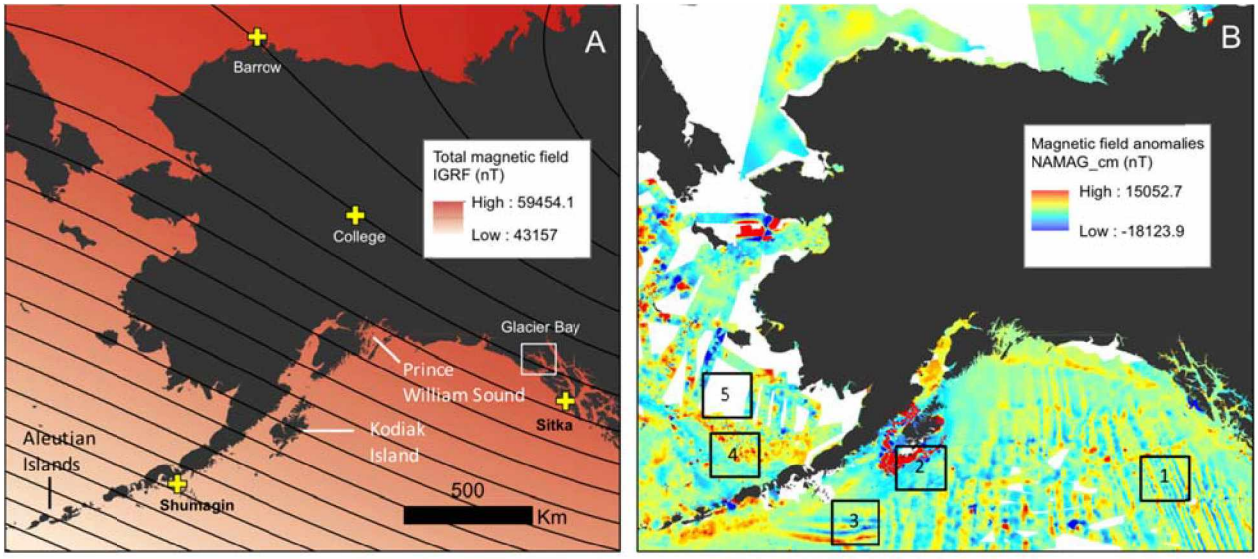


Figure 4.1. Magnetic field values in Alaska, USA. A) The main field at sea surface elevation modeled by International Geographic Reference Field (IGRF) increases from the southwest to the northeast (contour lines of 500 nT are shown). Four magnetic observatories in Alaska (see Figure 4.2) are indicated by yellow crosses. B) Magnetic field anomalies (red represents large positive anomalies, blue represents large negative anomalies, and green represents non-anomaly areas) occur throughout the region. Information on anomaly magnitude is available from a map with 1-km resolution by the North American Magnetic Anomaly Group (NAMAG). Numbered areas identify characteristics of the anomaly map referred to in the discussion of this chapter.

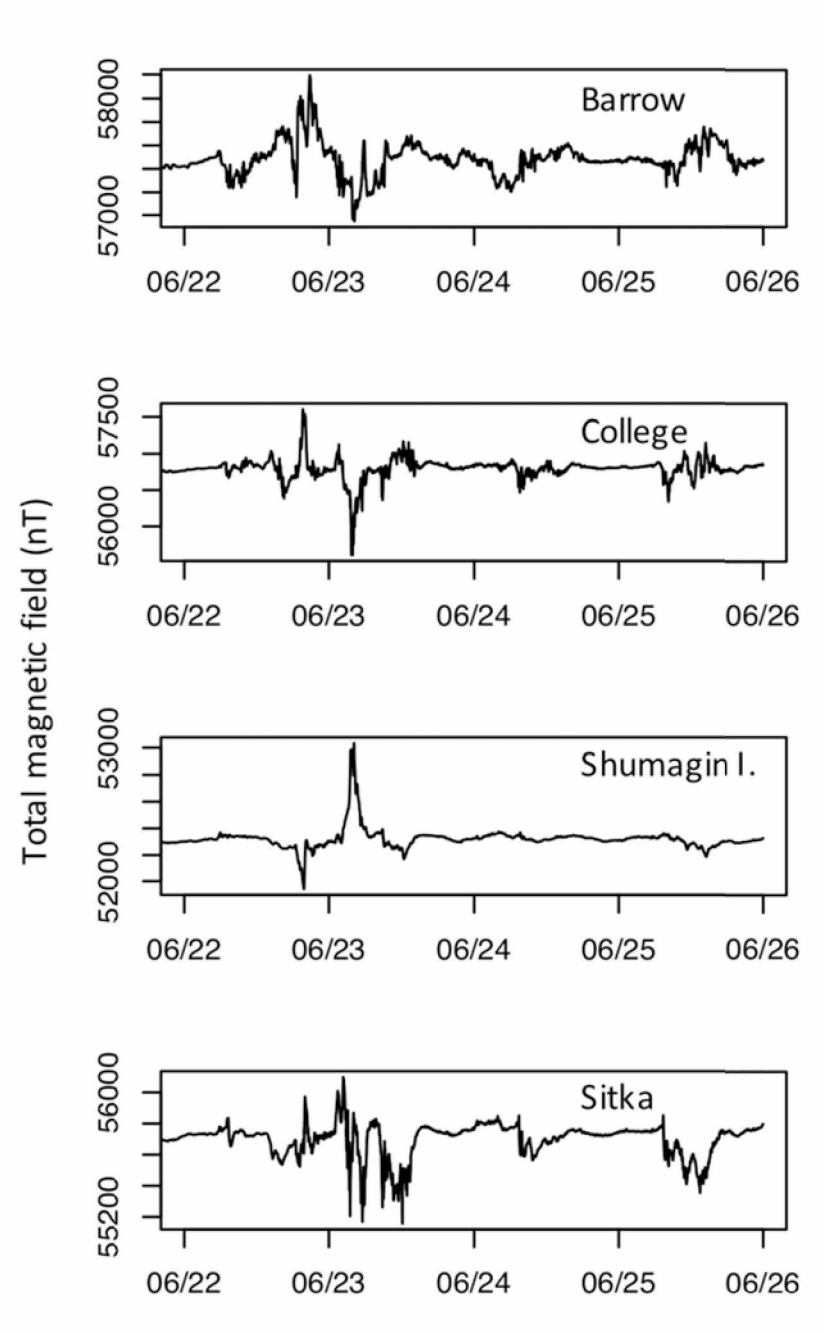


Figure 4.2. Example of solar storm effects on the total magnetic field at four observatories in Alaska, USA (locations shown in Figure 4.1), on June 22-26, 2015.

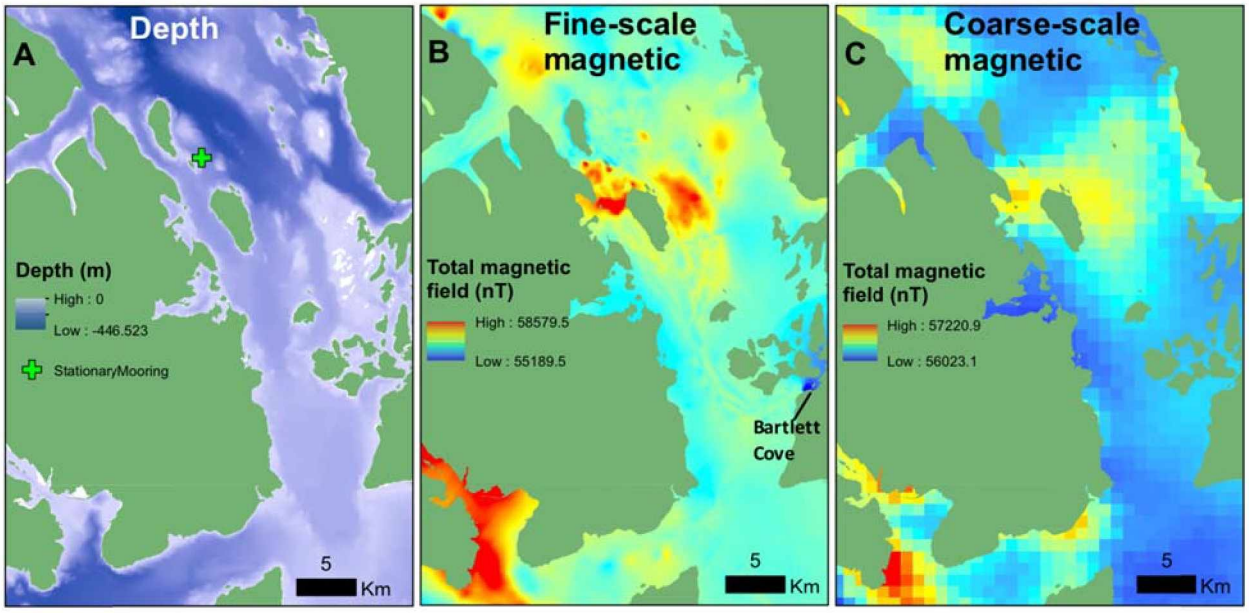


Figure 4.3. Maps of depth and magnetic field geolocation variables in the Glacier Bay National Park, Alaska, USA, study area. A) Depth (20-m resolution), B) fine-scale (100-m resolution) magnetic field, and C) coarse-scale (1-km resolution) magnetic field.

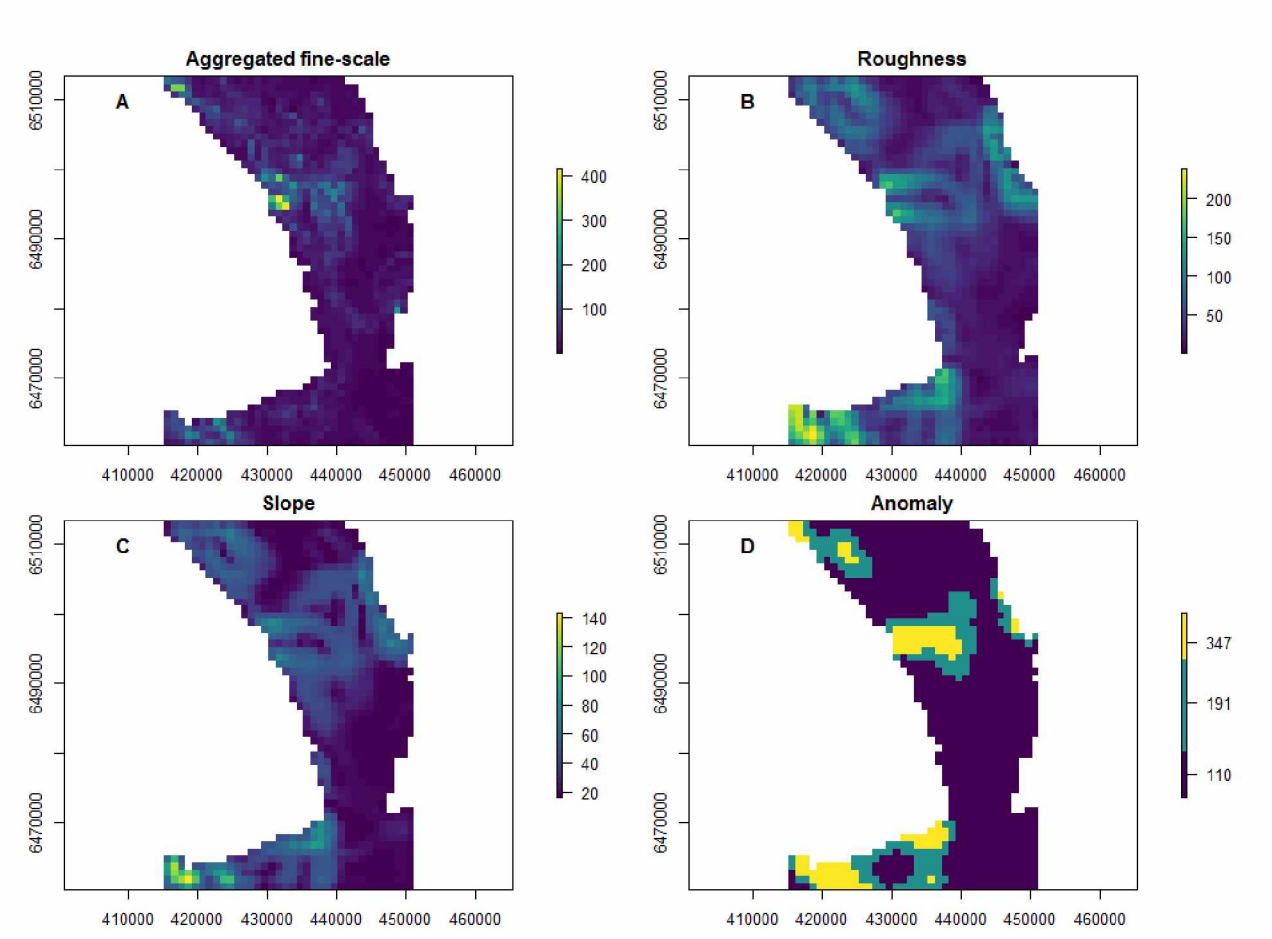


Figure 4.4. Four variance specification methods employed in geolocation of simulated fish movement trajectories. A) Standard deviation of fine-scale map (100 m) values within each1 km aggregated fine-scale model grid, B) the roughness method of obtaining variance (standard deviation of all adjacent cells for the coarse-scale map, C) the slope method of determining variance by relating fine-scale map values to the slope of the coarse-scale map, and D) variance specification based on the root mean square difference between measured and mapped values by magnitude of anomaly (see Figure 4.5). Note: islands and convoluted shorelines have been removed to simplify the simulation exercise.

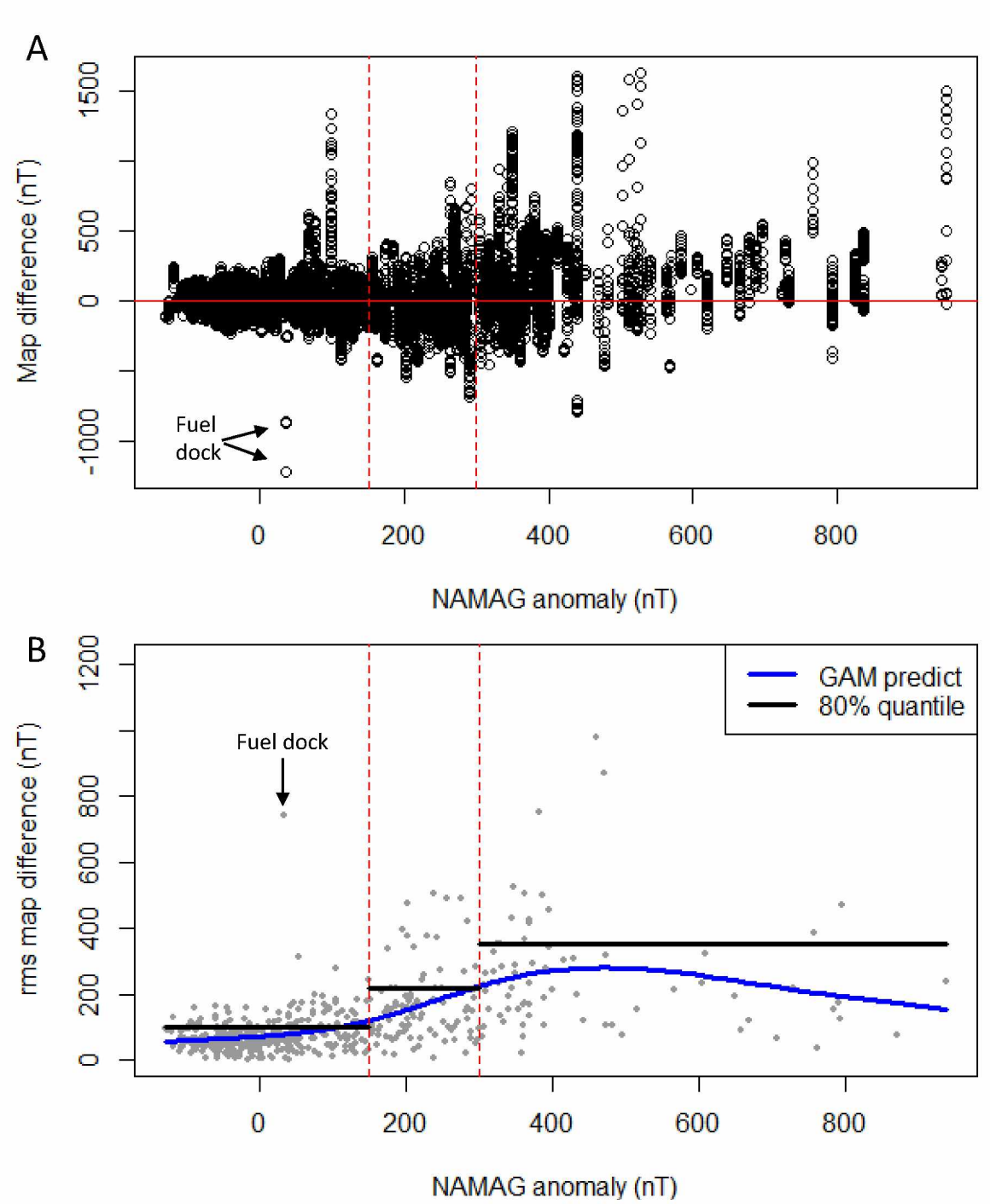


Figure 4.5. Differences between measured values from vessel survey of study area (see Supplement 4-1) and coarse-scale mapped values vs. NAMAG anomaly magnitude value for each coarse-scale grid cell. A) Raw differences between measured and mapped values. B) Root mean square (rms) differences between measured and mapped within each coarse-scale grid cell. Blue line indicates predicted values from the GAM of map difference vs. anomaly level. Red dashed vertical lines indicate division into three group based on NAMAG anomaly level: < 150 nT, 150 – 300 nT, and >300 nT. Black horizontal lines indicate 80% quantile values of rms difference within each of these bins used to specify variance used by anomaly level (see Figure 4.4D).

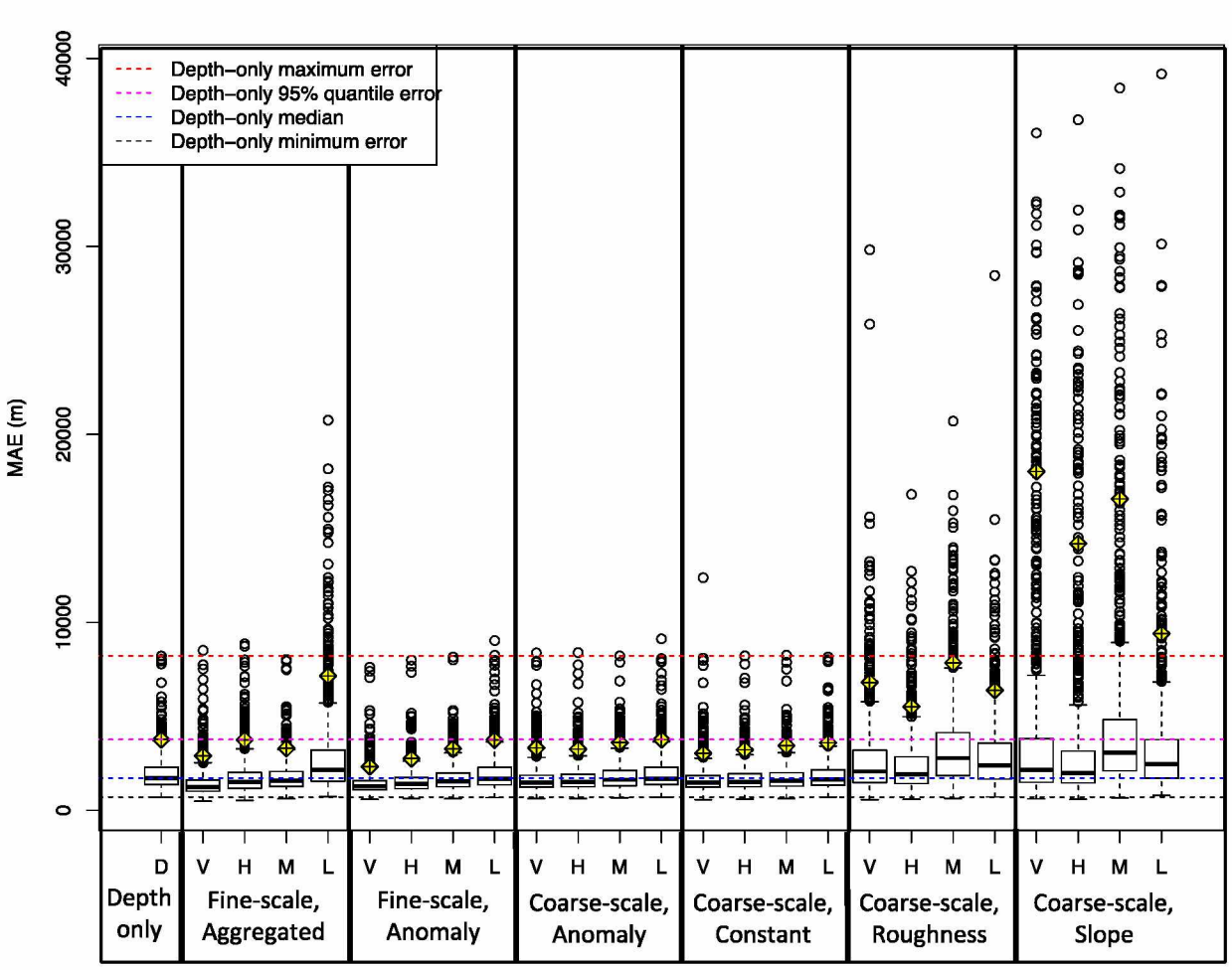


Figure 4.6. Performance of depth-only model (left) compared to 24 different magnetic and depth likelihood treatments. Higher performance is indicated by a lower mean absolute error (MAE) between estimated and known (simulated) locations for each of the 1000 trajectories. To assist in visual comparisons, horizontal dashed lines indicate the minimum (black), median (blue), 95% quantile (magenta), and maximum (red) MAE values observed for the depth-only treatment. Yellow diamonds indicate MAE 95% quantile values for each treatment. Treatment types are separated by thick vertical lines; within each treatment, V= very high, H = high, M= medium, and L= low tag measurement resolution. Fine-scale and coarse scale indicate which map was used for the treatment, and variance specification method is indicated by "aggregated" (Figure 4.4A), "anomaly" (Figure 4.4D), "constant" (165 nT for each grid cell), "roughness" (Figure 4.4B), and "slope" (Figure 4.4C).

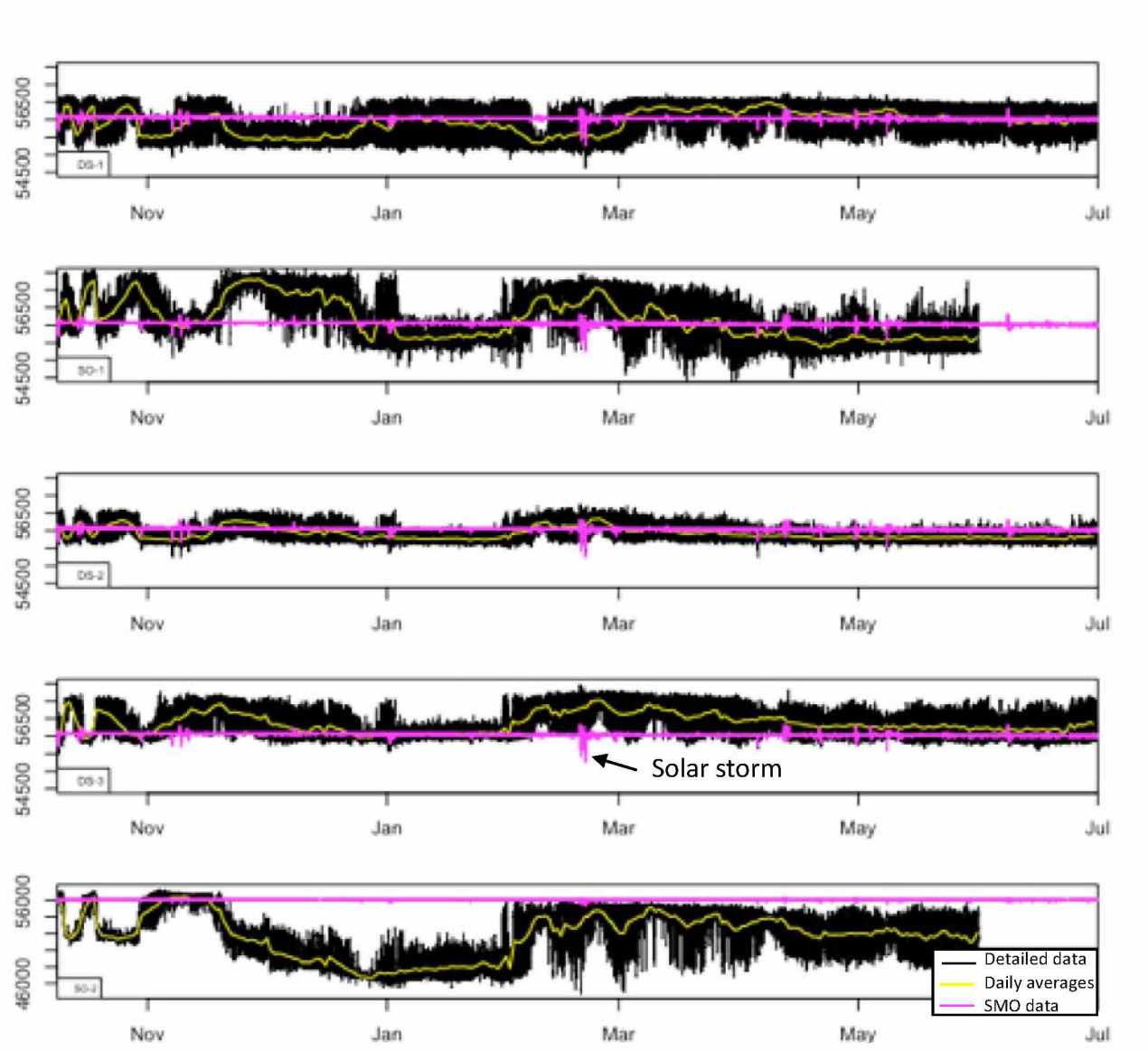


Figure 4.7. Magnetic field data provided by five archival tags from two manufacturers (DS = Desert Star, Marina, California, USA; SO = Star Oddi, Garōabær, Iceland) on a long-term stationary mooring in Glacier Bay National Park, Alaska, USA. Detailed magnetic field data are indicated by black lines and daily means by yellow lines. Total magnetic field data (1-minute resolution) from the nearby Sitka Magnetic Observatory (SMO) adjusted to the location of the mooring on day 1 of the deployment are indicated by the pink line. Note increased Y-axis values for SO-2 compared to other tags. Tags are plotted in the order they were attached to the mooring line (tag DS-1 was the shallowest tag and SO-2 the deepest).

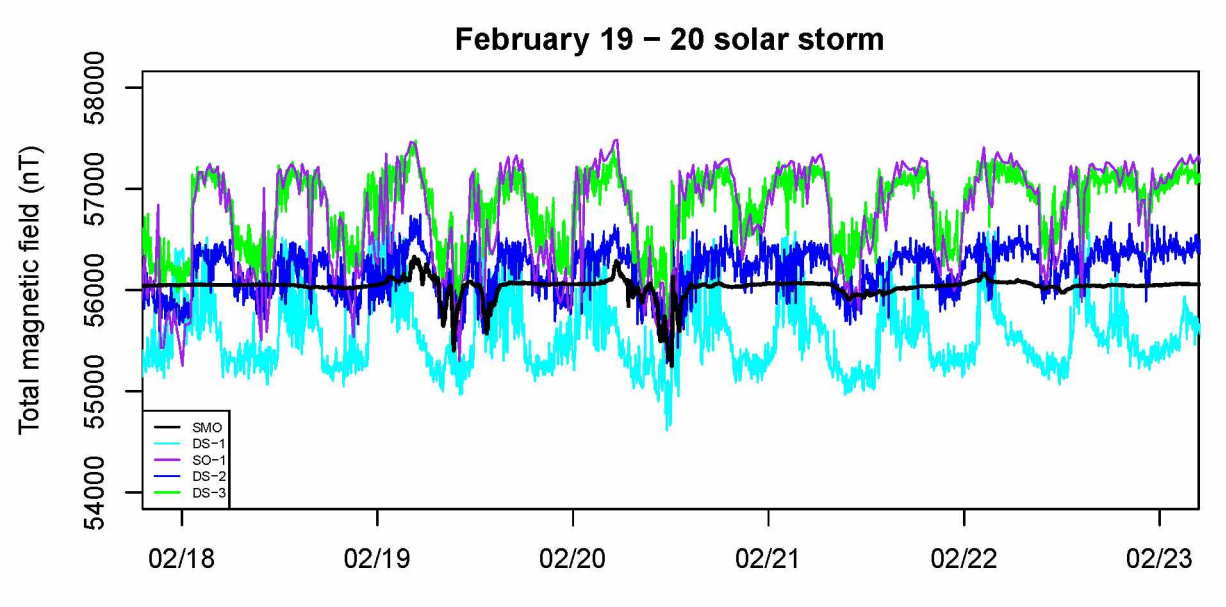


Figure 4.8. Total magnetic field values measured by the Sitka Magnetic Observatory (SMO, black line) during a solar storm compared to measurements from four stationary archival tags in Glacier Bay National Park, Alaska, USA.

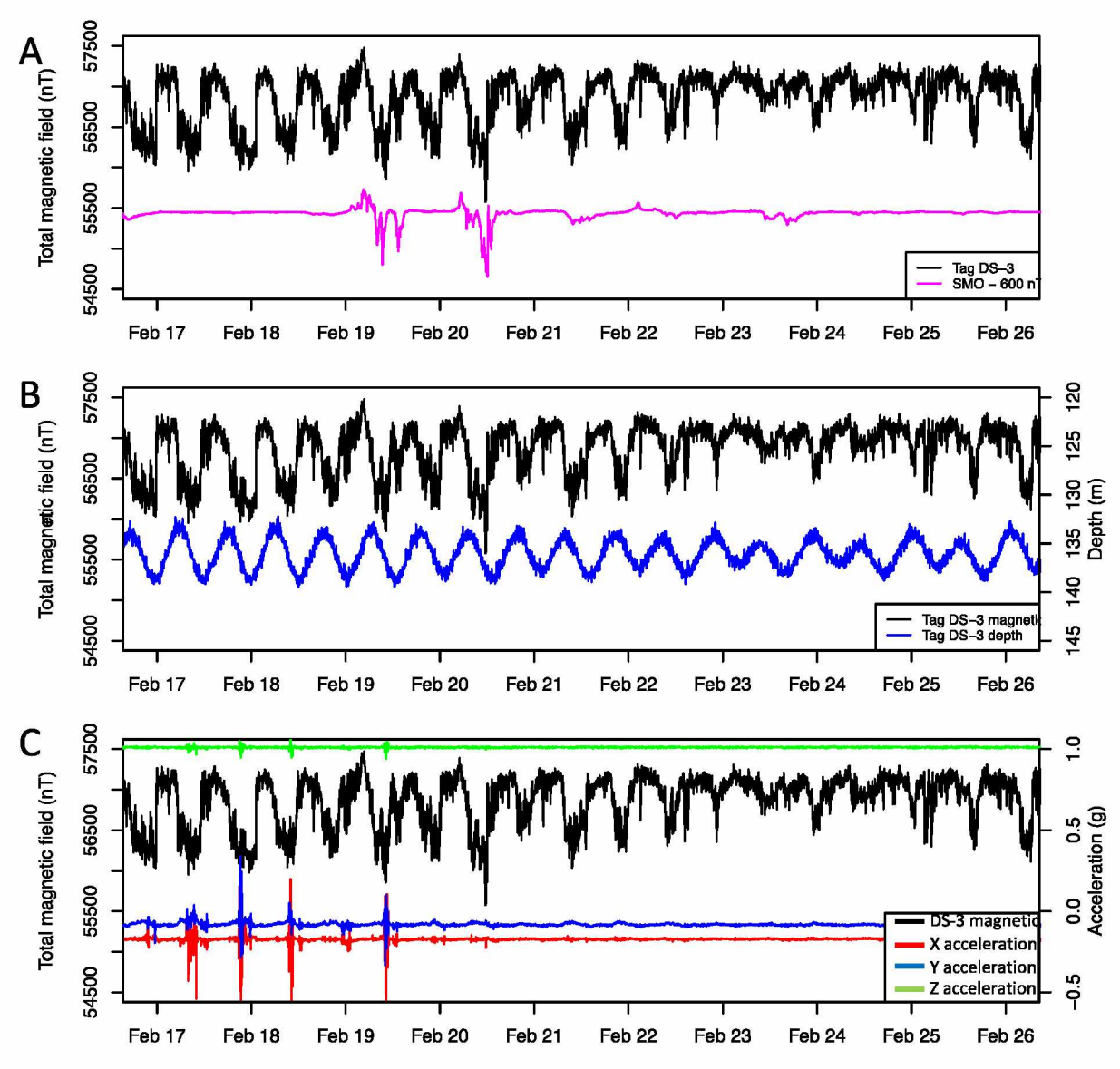


Figure 4.9. Variation in total magnetic field measurements recorded by archival tag DS-3 during a period of solar storm activity (magnitude 1000 nT) in Glacier Bay National Park, Alaska, USA. A) Detailed (4-minute resolution) total magnetic field measurements (black line) from tags compared to 1-minute resolution total magnetic field data from the Sitka Magnetic Observatory (pink line; offset by -600 nT to allow visual comparison). B) Detailed total magnetic field measurements (black line) from tags compared to depth (blue line) reflect a distinct tidal pattern in magnetic field data. C) Detailed total magnetic field measurements (black line) compared to tag orientation (tri-axial acceleration) over the time period, where orientation along the X axis is shown in red, Y axis in blue, and Z (vertical) axis in green.

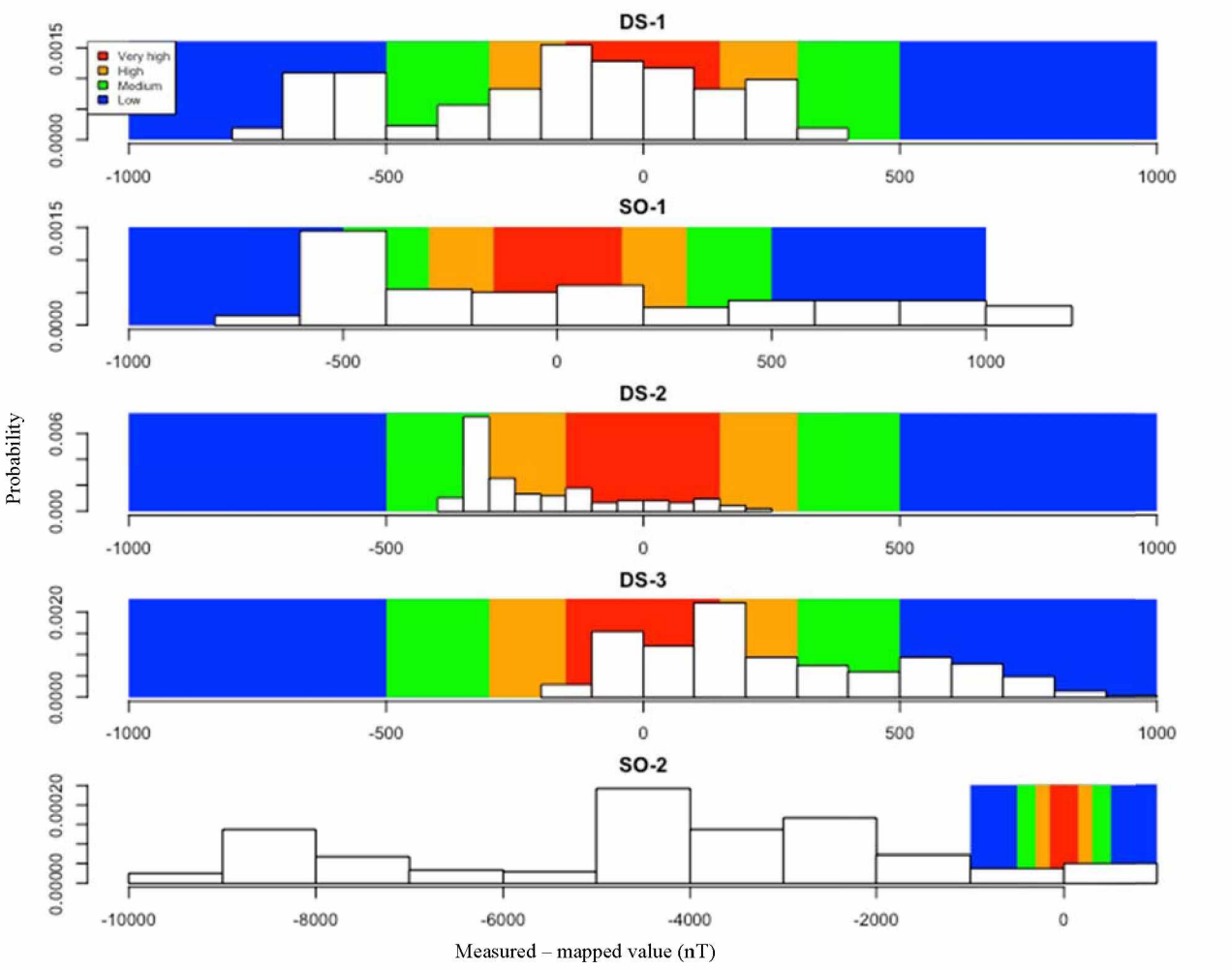


Figure 4.10. Histograms of the difference between the daily magnetic field measurements recorded by the five moored archival tags over the course of the 8-month deployment and the value of the fine-scale magnetic field map at the mooring location in Glacier Bay National Park, Alaska, USA. Colored polygons indicate range of tag resolutions used to define likelihoods for four levels of tag resolution in fish movement trajectories simulated via Hidden Markov Modeling. For example, to calculate the likelihood using the lowest tag resolution (blue), grid cell magnetic field probability density is integrated by limits of the daily measurement ± 1000 nT. Medium resolution (green) integrates the cell probability by ± 500 nT, high resolution (orange) by ± 300 nT, and very high resolution (red) by ± 150 nT. DS-1 would be considered a low-resolution tag because the histogram falls within the blue polygon, whereas DS-2 would be equivalent to a medium-resolution tag in the simulations because the histogram falls within the green polygon.

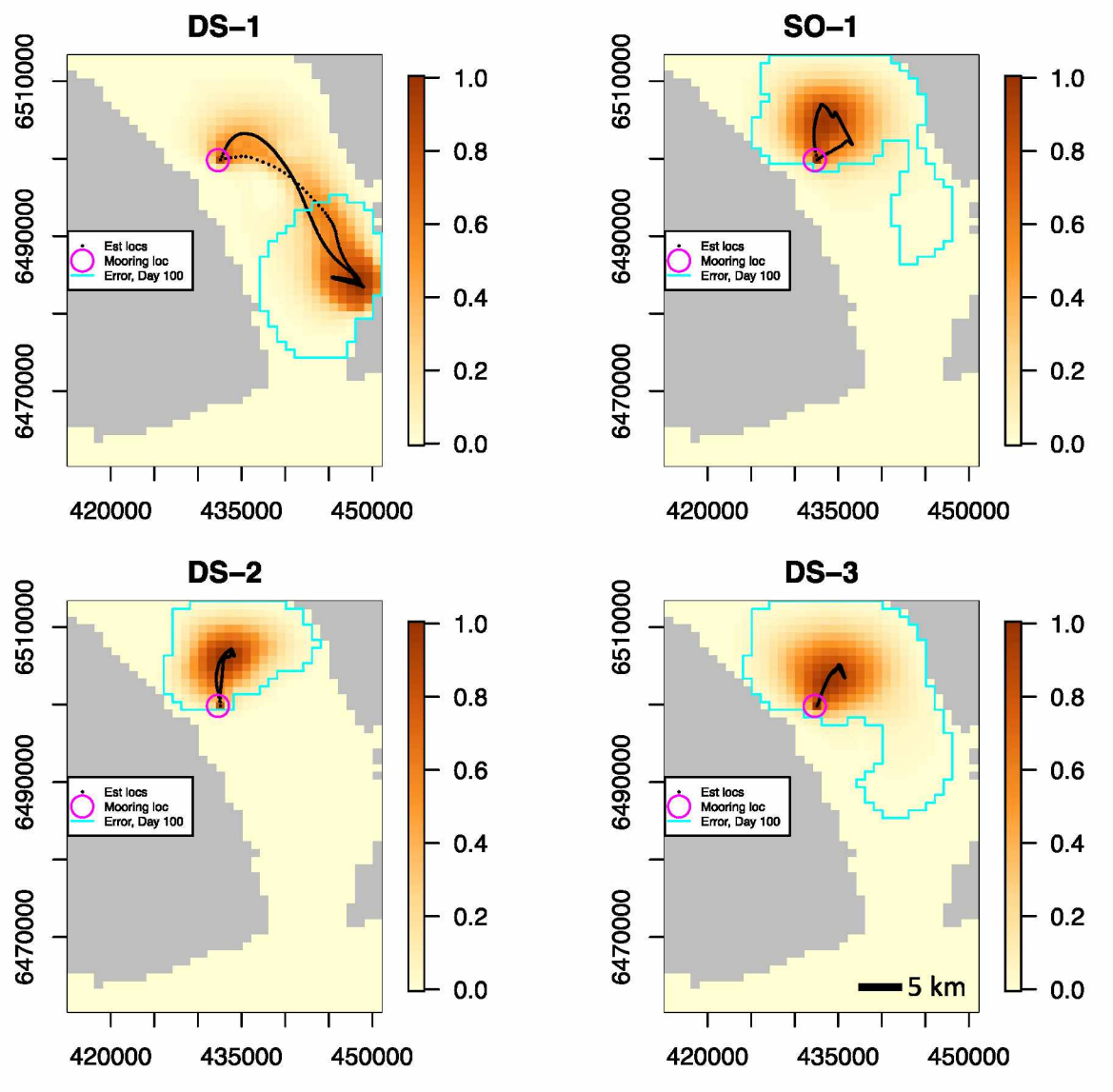


Figure 4.11. Hidden Markov Model (HMM) location estimates for 4 stationary geomagnetic tags based only on daily total magnetic field means. Colored surface indicates the residential distribution for the entire time period (October 9, 2013 – July 1, 2014); darker colors indicate higher probability; and black points indicate daily position estimates. The 99% probability polygon, which indicates estimated error associated with point locations, for Day 100 is shown in blue. The stationary mooring location is indicated by a pink circle.

4.9 Appendices

4.9.1 Supplement 4-1: Fine-scale magnetic field map of Glacier Bay

We created a 100-m grid of the total magnetic field in Glacier Bay, Alaska, USA, based on observations from a combination of fine-scale marine vessel and aerial survey data. We estimated the intensity of the magnetic field at locations visited during seven marine vessel surveys conducted over the course of one year (July 1, 2013 to July 1, 2014) using a mixed effects model as described below. To estimate magnetic field values for grid cells not measured during the vessel survey, we used the geostatistical technique of co-Kriging to leverage fine-scale aerial survey data available for the entire study area. This resulted in a fine-scale map that covers all of Glacier Bay and the waters extending into Icy Strait and into the Gulf of Alaska (Figure 4.3B).

A GEM (Markham, Ontario, Canada) GSM-17 Overhauser magnetometer/gradiometer was attached to the bow of an aluminum vessel (Figure S4.1-1). The instrument recorded magnetic field readings (resolution 0.01 nT) and GPS coordinates every second. The two sensors were mounted on a vertical bar and placed 0.5 m apart to sense gradients in the vertical component of the magnetic field. Because the vessel was aluminum, and the sensors were located far enough from electronic components to prevent interference, the change in magnetic field with change in direction of the vessel was negligible (< 50 nT). The magnetometer was deployed during 7 tracking trips; 5 during the summer of 2013 and 2 during the summer of 2014. Vessel speeds were generally < 7 knots when magnetic data were collected. After spurious records were removed, this generated a data set of 881,292 observations collected over a time period of July 1, 2013 to July 1, 2014 (Figure S4.1-2).

Stationary test tags and magnetic observatory data were used to control for temporal change in magnetic field strength during the vessel survey. While the magnetometer on board the vessel was recording data, a stationary archival tag (SeaTagMOD, Desert Star, Marina, California, USA) was placed on shore to monitor changes in the magnetic field due to solar activity. The tag was put on ice in a darkened cooler so that it would maintain a constant temperature. If the tags are not exposed to light, changing temperature, or a change in orientation, the tag resolution in the measurement of the total

magnetic field has a range of approximately ± 100 nT (Supplement 4-2). In addition, magnetic field values from the Sitka Magnetic Observatory (SMO, approximately 100 miles from Glacier Bay) were available each minute as an additional source of information on the magnitude of change in the magnetic field over time during the surveys. Although the absolute value of SMO differed from the stationary archival tag measurements, changes were of similar magnitude (Figure S4.1-3). A good agreement between stationary test tags and the SMO data, albeit with a slight lag at times, suggests that the detailed data recorded by the SMO provided a reasonable indication of temporal change in the magnetic field in the study area.

We aggregated the magnetic field observations into 100 m x 100 m cells that reflected the range of spatial autocorrelation present in the vessel data. First, we removed the annual temporal trend in the main field (-85.147 nT/year, $p < 2e\text{-}16$) using a linear regression on all data available from SMO over the course of the year (Figure S4.1-2). Then we divided the observations into segments that corresponded to each transit through each 100 m^2 grid cell on each trip to account for potential differences in the magnetic field measurements due to the direction of vessel travel and small-scale temporal magnetic fluctuations. We fit a hierarchical linear model for observations vs. magnetic field gradient, using trip ID and transit ID as grouping variables. The intercept was the "population" estimate of the magnetic field in each cell. If just one observation was available, that value was recorded as the cell value. Mapped values were converted to anomaly data by subtracting International Geomagnetic Reference Field (IGRF) values for July 1, 2013. All analyses were conducted using the R statistical program (R Core Team, 2017).

Total magnetic field values collected from the vessel survey revealed several distinct anomaly areas in the study area (Figure S4.1-4A). The standard deviation of aggregated magnetic field observations in each 100 m grid cell was usually less than 150 nT (Figure S4.1-4B). The primary purpose of the survey trips was to track Pacific halibut (*Hippoglossus stenolepis*) with acoustic telemetry, so most time was spent in the vicinity of tagged fish that tended to establish home ranges (Nielsen and Seitz, 2017). As a result, some locations received many more observations than others (Figure S4.1-4C).

Aerial survey data for the study area (Figure S4.1-5) were used to estimate magnetic field values for grid cells not visited during the marine vessel survey. Aerial data for Glacier Bay were collected in 1976 (Brew et al., 1978; Connard et al., 1999) using a modified ASQ-10 fluxgate magnetometer (precision 0.1 nT). Magnetic field observations were 200 m apart on flight transects spaced 1600 m apart. Magnetic field measurements were adjusted to an elevation of 305 m above sea level (Brew et al., 1978; Connard et al., 1999).

Because the marine vessel collected magnetic field data at sea level while the aerial survey data provided magnetic field values for an elevation of 305 m, values were assumed to vary similarly in space though the absolute magnitude would not be the same for the two data sets. We used the geostatistical technique of co-Kriging to apply information about the spatial variation observed in the aerial survey while preserving the magnitudes from the vessel-based data. Co-kriging was conducted using the Geostatistical Analyst extension in ArcGIS 10.3.1 for Desktop. Following co-Kriging, IGRF values for July 1, 2013 were added to the combined anomaly map to represent total magnetic field values on that day.

4.9.1.1 Literature cited

Brew, D. A., Johnson, B. R., Grybeck, D., Griscom, A., Barnes, D. F., Kimball, A. L., Still, J. C., et al. 1978. Mineral resources of the Glacier Bay National Monument wilderness study area, Alaska. USGS Open-file Report 78-494.

Connard, G. G., Saltus, R. W., Hill, P. L., Carlson, L., and Milicevic, B. 1999. Alaska digital aeromagnetic database description. Open-file Report 99-0503.

Nielsen, J. K., and Seitz, A. C. 2017. Interannual site fidelity of Pacific halibut: potential utility of protected areas for management of a migratory demersal fish. ICES J Mar Sci, 74: 2120–2134.

R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Figure S4.1-1. Configuration of magnetometer mounted to the bow of an aluminum vessel that conducted surveys in Glacier Bay National Park, Alaska. Horizontal bar for deploying hydrophones was also aluminum. Two vertically mounted sensors on the magnetometer detected magnetic field magnitude and gradients.

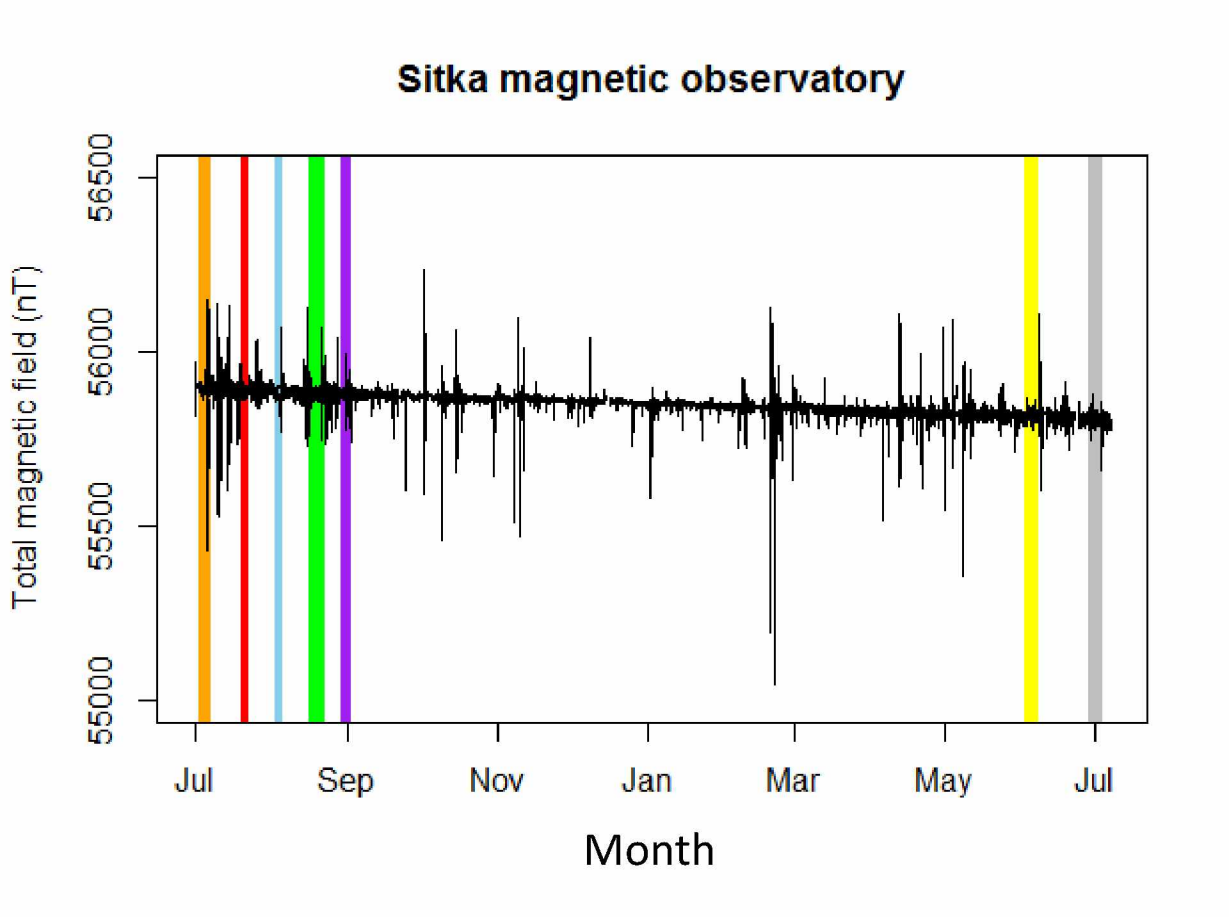


Figure S4.1-2. Temporal fluctuations in the magnetic field over the course of the study in Glacier Bay National Park, Alaska during 2013 and 2014. Measurements from the nearby Sitka Magnetic Observatory, Alaska (black line) indicate the effects of solar storms (spikes) and slow change in the main field (linear decline over time). Colored bars indicate times when magnetic data were collected for mapping (Trip 1 = orange, Trip 2 = red, Trip 3 = blue, Trip 4 = green, Trip 5 = purple, Trip 6 = yellow, and Trip 7 = gray).

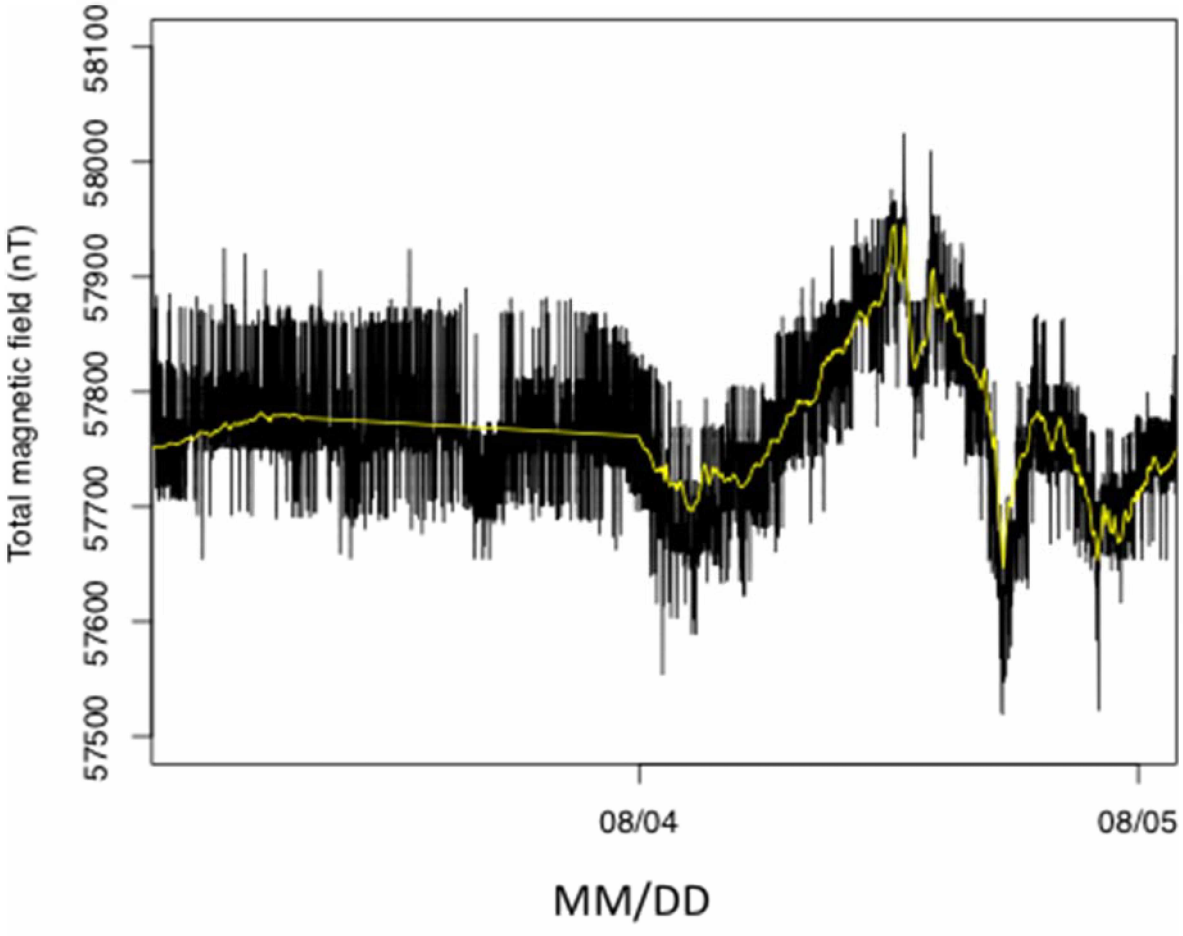


Figure S4.1-3. Temporal fluctuation in the total magnetic field during survey Trip 3 in Glacier Bay National Park, Alaska. Total magnetic field measurements from a stationary Desert Star archival tag adjusted for temperature at 30 second intervals (black line) and simultaneous magnetic field measurements from the Sitka Magnetic Observatory (SMO) at one-minute intervals (yellow line). SMO values were adjusted by +1880 nT to match the tag baseline.

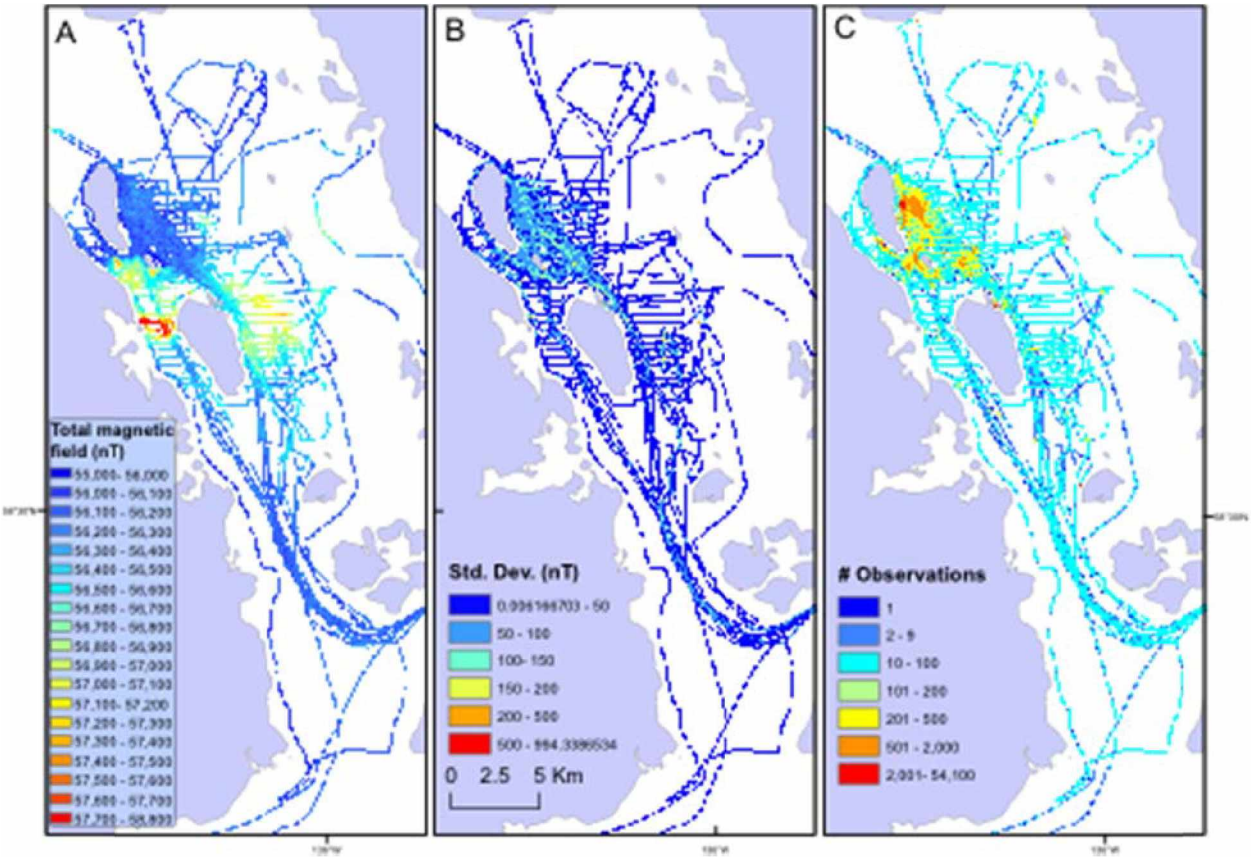


Figure S4.1-4. Magnetic field map data collected by vessel during seven surveys over the course of one year in Glacier Bay National Park, Alaska. Total magnetic field values (A) and standard deviation of aggregated values (B) for each 100 m grid cell. C) Observations were clustered in areas where tagged fish were being tracked with acoustic telemetry (the primary purpose of the surveys; Nielsen and Seitz 2017) with vessel transects conducted in regions adjacent to areas where tagged fish resided (i.e., grid cells with number of observations > 200).

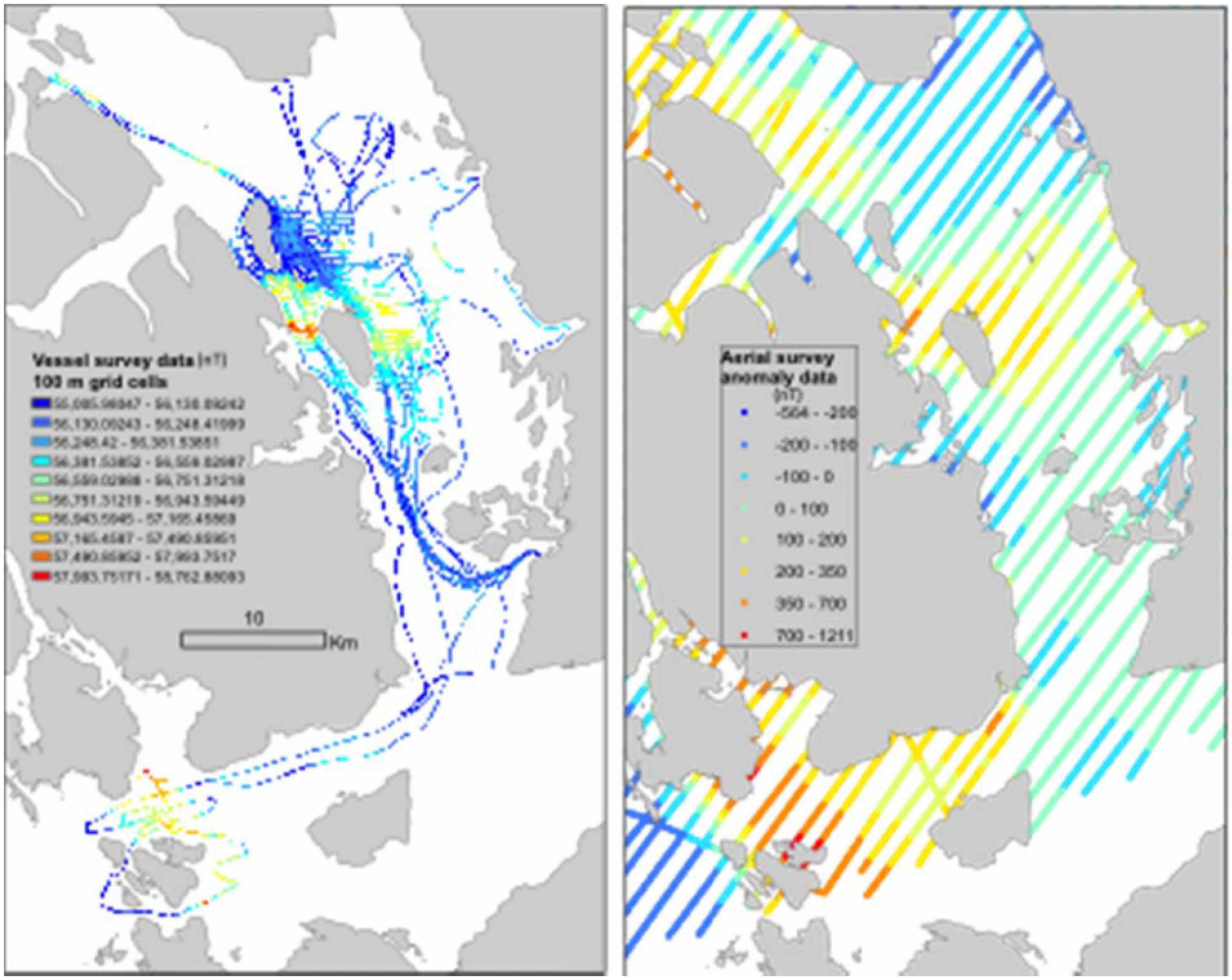


Figure S4.1-5. Vessel survey total magnetic field map (100 m resolution), left, and aerial survey magnetic anomaly observations for Glacier Bay National Park, Alaska (Brew et al., 1978; Connard et al., 1999), right.

4.9.2 Supplement 4-2: Geomagnetic tag measurement artifacts

Magnetic field measurements recorded by geomagnetic archival tags do not always represent the actual value of the Earth's magnetic field at a given location and time because they can be influenced by tag conditions such as orientation or temperature. These issues related to sensor and tag technology can produce artifacts in archived data sets that mask actual changes in the magnetic field or produce misleading patterns that seem like horizontal displacement when they are really not. In this appendix, we discuss two common tag measurement artifacts in the context of the temporal patterns observed for the five archival tags that were deployed on the long-term stationary mooring (Figure 4.7). As stated in the caveats section of the chapter, the tag data presented here were obtained from the earliest versions of geomagnetic tags (2012 – 2013) supplied from Desert Star (Marina, California, USA) and Star Oddi (Garōabær, Iceland). Improvements have likely been made since then, but researchers should ask the manufacturers how they have addressed the following tag issues and confirm that they have been addressed prior to deployment on fish.

4.9.2.1 Orientation

Magnetic sensors on board archival tags are subject to an array of effects that can influence the values that they measure when held in different orientations (Yang et al., 2017). These effects include hard and soft iron effects caused by other nearby tag components such as batteries. Hard iron effects result from constant distortion of the magnetic field in the vicinity of the sensor from materials such as iron that have their own magnetic field. Soft iron effects are caused by materials that are susceptible to magnetization when exposed to a magnetic field; therefore, the distortion caused by these materials increases when aligned with external fields (such as the Earth's magnetic field). In addition, rotational errors can be caused by non-orthogonal sensors and mis-alignment of sensors with the instrument.

Such errors can be demonstrated simply by placing the tag in different orientations and examining the total magnetic field readings measured by the tags at each orientation (Figure S4.2-1). In the example shown, the tag has a standard deviation of 43 nT around the mean measured value (range ± 100 nT) when orientation, temperature, and light intensity are constant. However, the measured value of the magnetic

field varies from the true value at the tag location by more than 2000 nT depending solely upon the tag orientation.

The measurement issues related to change in tag orientation produces patterns in the total magnetic field measured by stationary tags attached to mooring lines over long time periods. We first observed these patterns from a Desert Star SeaTagMOD attached to a long-term stationary mooring near Kodiak, Alaska, USA, previous to conducting the current study (Figure S4.2-2). The tag was mounted loosely to the mooring line so that in the case that the mooring line could not be recovered, the tag could still pop off and provide summarized data. Therefore, the motion of the tag was constrained, but it was not fixed tightly to the mooring line. The range in detailed data was 3400 nT over the course of the deployment, and the range of daily means fluctuated within a range of 1630 nT. When magnetic field observations were restricted to only one tag orientation, the range in detailed measurements was reduced to 1280 nT and the range in daily means was reduced to 819 nT.

In contrast to the Desert Star SeaTagMOD on the Kodiak mooring, the stationary tags in the current Glacier Bay study were tightly affixed to the line, so that changes in tag orientation were related to change in orientation of the mooring line itself. When we restricted tag magnetic field measurements from the Glacier Bay stationary tags to only one tag orientation based on acceleration measurements (Table S4.2-1), the range in detailed magnetic field measurements was not reduced to the degree observed for the Kodiak tags (Figure S4.2-3). In addition, restricting the tag orientation did nothing to reduce the "wandering" pattern, where measurements fluctuated between positive and negative values relative to the known value at the mooring location. Therefore, either very small changes in orientation cause large effects on magnetic field measurement, or some other aspect of tag movement on the mooring line or internal architecture may be causing the "wandering pattern" observed in the magnetic field data. However, because the magnetic field sensors and archival tags are complicated, we leave further discussion of possible explanations for the temporal patterns observed on mooring lines to engineers and others more familiar with the details of their operation.

4.9.2.2 Temperature

All geomagnetic archival tags are sensitive to temperature change due to the nature of the sensors employed. In general, as temperature increased, the reported value of the magnetic field decreased for both tag models that were deployed (Figure S4.2-4). However, differences in change in measured magnetic field per unit of temperature change may vary by tag.

4.9.2.3 Literature cited

Yang, D., You, Z., Li, B., Duan, W., and Yuan, B. 2017. Complete tri-axis magnetometer calibration with a gyro auxiliary. Sensors (Basel, Switzerland), 17: 1223.

4.9.2.4 Acknowledgements

We thank Katy Echave (NOAA, Ted Stevens Marine Institute) for providing the Desert Star SeaTagMOD deployed on the Kodiak mooring buoy. We thank the International Pacific Halibut Commission and the F/V Kema Sue for deploying and recovering the Kodiak mooring assembly.

Table S4.2-1. To reduce effects of rotational distortion on magnetic field measurement, archival tag data were restricted to only one orientation to determine whether magnetic field values from the subset data still exhibited temporal patterns unrelated to known temporal fluctuations in the field. To restrict measurements to one orientation, accelerometer X, Y, and Z axes were each constrained to a specific range of values. These are measured in units of gravity (g) for Desert Star (Marina, California, USA) SeaTagMOD (DS) tags and degrees (d) for Star Oddi (Garōabær, Iceland) DSTmagnetic (SO) tags. The number of records comprising the subset data and percent of the full data set are also shown. Subset data are shown in Figure S4.2-3.

TagID	X-axis range	Y-axis range	Z-axis range	# records	% original
DS-1	> -0.175, < -0.17 g	> 0.034, < 0.036 g	> 0.965, < 0.97 g	4767	5.0
DS-2	> 0.083, < 0.09 g	> -0.09, < -0.08 g	> 1.0, < 1.03 g	17482	18.3
DS-3	> 0.08, < 0.09 g	> -0.09, < -0.08 g	>1.005, < 1.013 g	17642	18.5
SO-1	> -11.2, < -10.8 d	> 10.4, < 10.8 d	> 68.3, < 68.6 d	3264	19.3
SO-2	> -4.9, < -4.5 d	> 0.25, < 0.6 d	> 84.8, < 85.2 d	1986	11.8

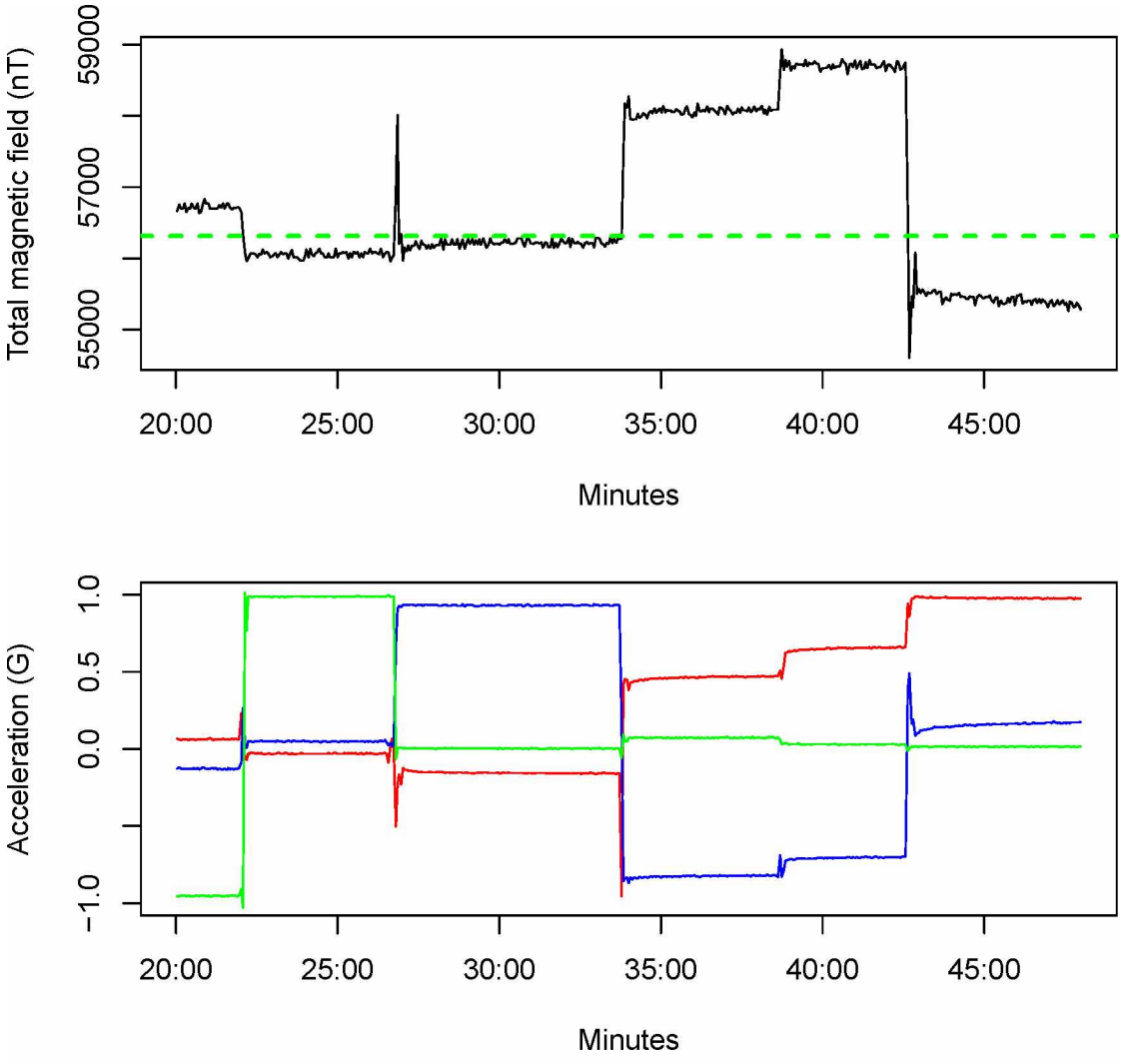


Figure S4.2-1. Effect of tag rotation on total magnetic field values measured by a Desert Star (Marina, California, USA) SeaTagMOD in Juneau, Alaska, USA. Top: Magnetic field values collected during a 30-minute period of low light, when temperatures were 2 – 2.5°C. Dashed line indicates known value at tag location determined with a G-857 proton precession magnetometer (Geometrics, San Jose, CA, USA). Bottom: Acceleration values indicate different tag orientations during the same time period (when a tag is not in motion accelerometers measure sensor angle relative to the direction of gravity). Acceleration X axis is shown in red, the Y axis in blue, and the Z axis in green.

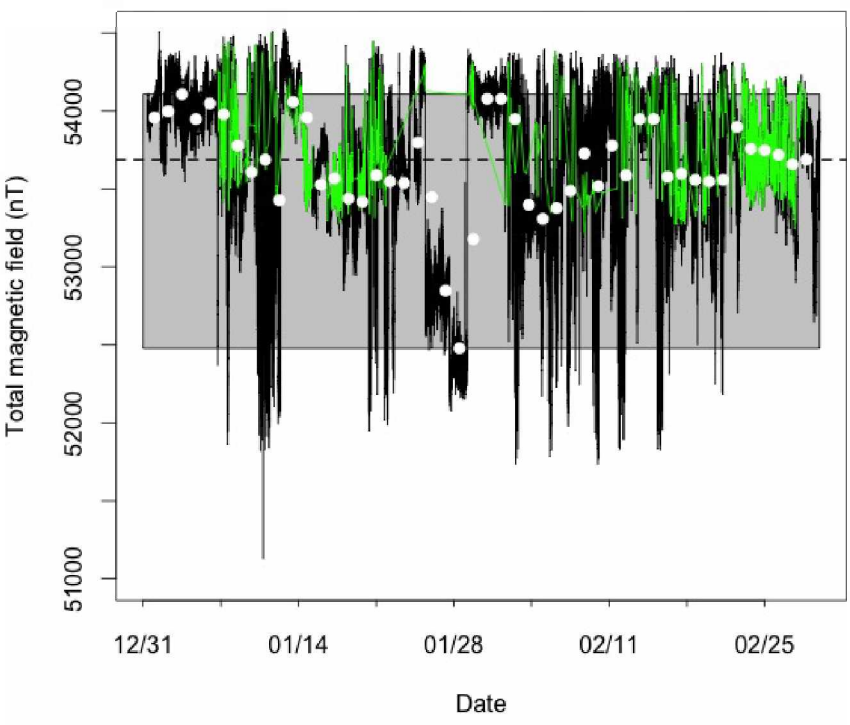


Figure S4.2-2. Total magnetic field values from a 3-month deployment of a SeaTagMOD geomagnetic archival tag (Desert Star, Marina, California, USA) loosely affixed to a mooring line near Kodiak, Alaska, USA. Detailed (4-minute resolution) magnetic field measurements are shown in black, and daily means are indicated by white circles. The dashed line indicates the mapped value at the mooring location and the gray box indicates the range in daily averages. A subset of detailed measurements where the tag is restricted to only one tag orientation is indicated by green lines.

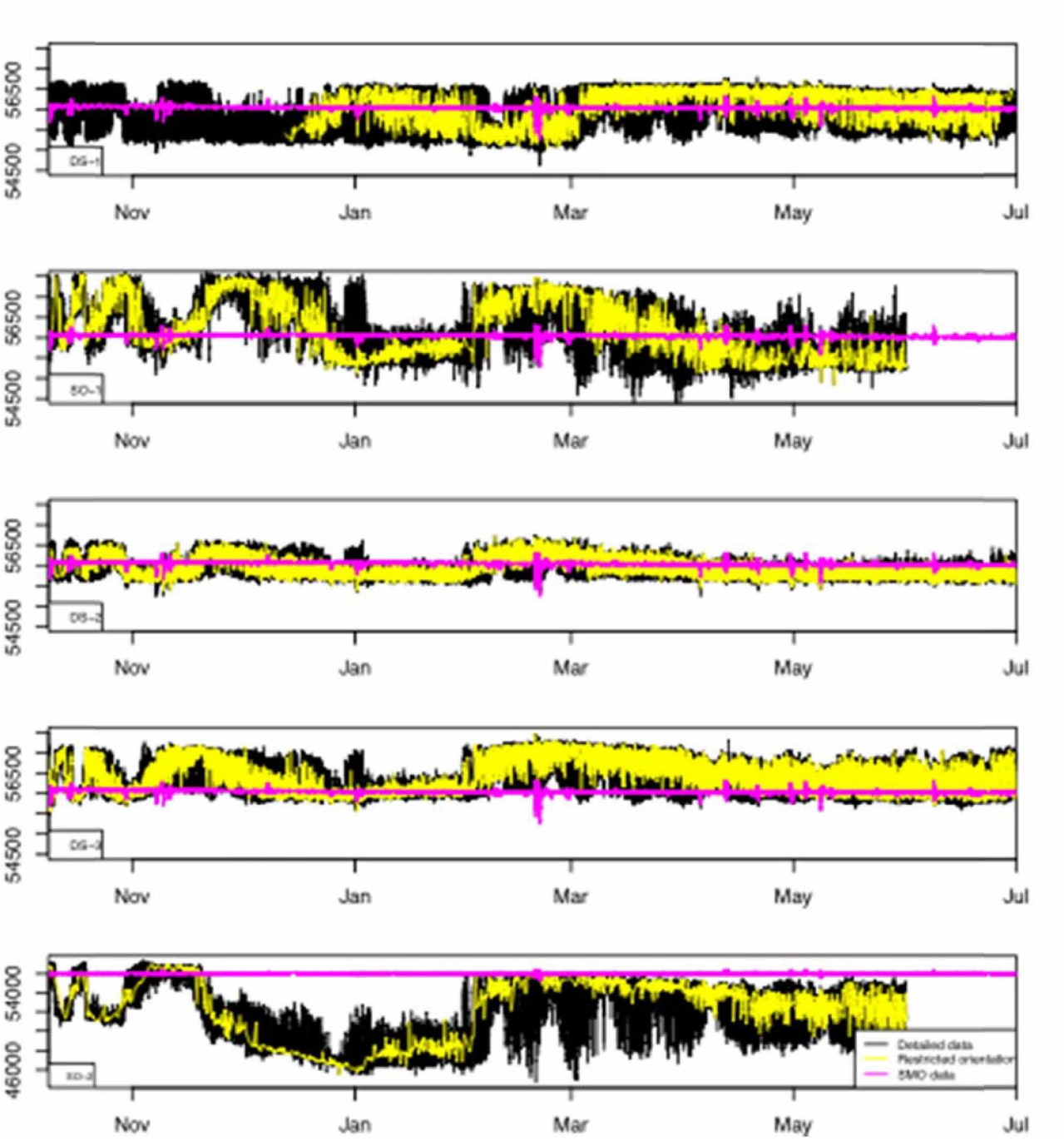


Figure S4.2-3. Detailed total magnetic field strength measurements from the five stationary archival tags deployed in Glacier Bay, USA, on the long-term mooring (also shown in Figure 4.7 of the main text; black lines). Measurements when tag was restricted to only one orientation (Table B-1) are indicated by yellow lines. The change in magnetic field values recorded by the near-by Sitka Magnetic Observatory (SMO) is indicated by pink lines.

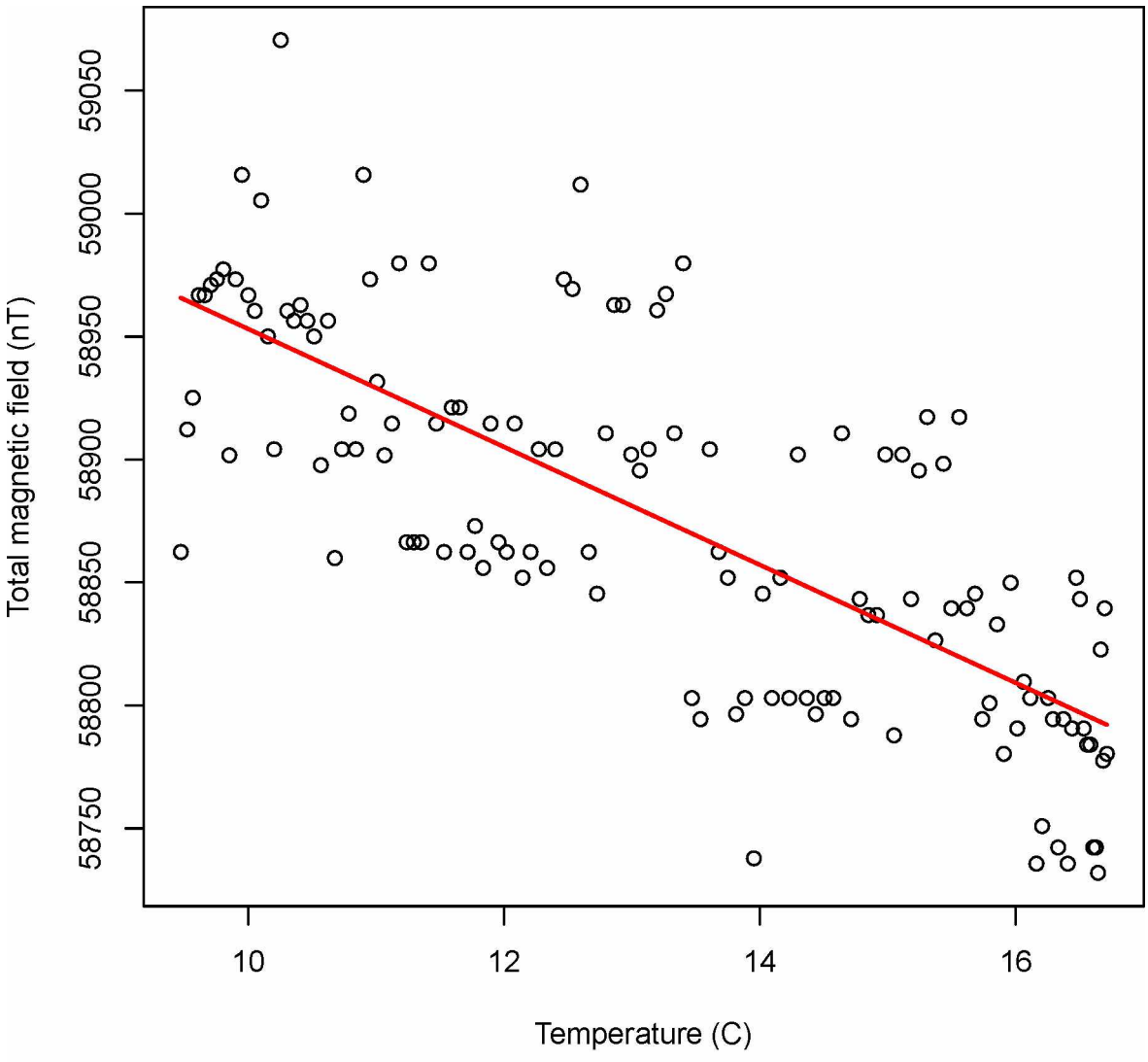


Figure S4.2-4. Relationship between change in temperature and change in measured magnetic field values for a Desert Star (Marina, California, USA) SeaTagMOD in stationary and low light conditions in Juneau, Alaska. Red line indicates linear model fitted slope of -24.0 nT per °C ($p<2e16$).

General conclusion

This dissertation has presented new methods for obtaining detailed information on the movement of demersal fish at multiple scales in Alaska. It has also provided insights into behaviors of Pacific halibut, such as seasonal and annual site fidelity, that have important implications for management such as spatial and temporal closures of preferred habitat. I conclude the dissertation by discussing the benefits of deploying multiple types of electronic tags, the importance of the Hidden Markov Model (HMM) for fish geolocation, and examples of ways in which the detailed information provided by these methods may be important for fisheries and marine resource management.

One of the unexpected results of my research was the value of double-tagging fish with archival and acoustic tags. The ability to detect and quantify the spatial scale of home range behavior during summer foraging with acoustic telemetry combined with the knowledge of winter migration timing from PSAT data provided valuable insights into the unexpectedly small spatial scales of site fidelity and homing that at least some adult Pacific halibut are capable of demonstrating. The spatial scales of summer home ranges and ability to navigate back to them were similar to some *Sebastes* rockfish species (Mitamura et al., 2012). The scale and nature of site fidelity of Pacific halibut to specific locations suggests that characterizing attributes of these locations may provide insights into the spatial distribution of adult Pacific halibut during summer. However, future research on the spatial scales of site fidelity in different habitat types and for different size classes is needed to interpret the results observed for Pacific halibut in Glacier Bay National Park, Alaska in the context of the overall population.

In addition to increased precision in knowledge of annual movement patterns, double-tagging with both archival and acoustic tags provided insights into movement patterns that were greater than would be achieved by each type of tagging technology alone. For example, based on archival tagging records, some fish with home range behavior were inactive during the day when they were acoustically tracked, but were more mobile at night, particularly in deeper waters (Chapter 2). Since fish were only tracked during the day, home range estimates for fish that were more active at night were likely biased toward a smaller scale of movement than was actually exhibited over the course of a 24-hour period. On

the other hand, the observations of home range behavior from acoustic telemetry helped to identify characteristic patterns of home range behavior in archival tag data sets that provided inference into the duration of home range behavior after tracking trips were discontinued for the summer. These insights from double-tagging were particularly important for interpreting the results of manual tracking, for which the frequency and spatial scale of the tracking trips determines the potential for detection of tagged animals.

Finally, double tagging fish can improve HMM accuracy as acoustic tag locations can be directly incorporated into HMM data likelihoods and can be used for independent validation of HMM model estimates. Double-tagging with acoustic and archival tags has also been conducted for Atlantic salmon (Strøm et al., 2017) and Atlantic cod (Liu et al., 2017), where acoustic telemetry positions were used to enhance and/or validate HMM results. Given the continued need for validation of HMM model estimates (see below), the practice of double-tagging will be valuable for understanding prediction error and potential biases of the HMM for different geolocation applications.

Another important outcome of my research has been insights into procedures for adapting the HMM developed for geolocation of Atlantic cod (Pedersen et al., 2008; Thygesen et al., 2009) for demersal fishes in Alaska. Adapting the HMM requires changes to the data likelihood model based on the type of available geolocation data (e.g., harvest-recovered or PSAT), characteristics of the study area (e.g., orientation and strength of gradients), and behavior of the species (e.g., pelagic or demersal) for each application. A detailed description of the HMM and depth-based data likelihood models for demersal fish in Alaska is provided as an appendix to this dissertation. The components of the HMM are fairly simple: a gridded study area, a movement model based on diffusion, and a customized data likelihood model that provides the likelihood that the fish occupied each grid cell based on the value measured by the tag at each timestep. HMM studies usually provide information about the data likelihood model, the grid size used, and the value of diffusion chosen. However, aside from reporting chosen values, there is little guidance about how to choose those parameters for a particular application in the literature (but see Braun et al., 2017, Appendix S1). The results of Chapter 3 (depth-based likelihood) suggest that consideration of

study area attributes such as bathymetry heterogeneity is important for the choice of both grid size and data likelihood specification method. The results of Chapter 4 (depth- and geomagnetic-based likelihoods) suggest that, when combining multiple geolocation variables, sub-optimal variance specification methods for one variable can reduce the model performance compared to a single (more precise and accurately mapped) variable modeling approach. Therefore, the ability to customize these fixed parameters for different applications, and individual geolocation variables, could be important for model performance. These observations are significant because new tools for implementing the model are available that may not allow optimal customization by application. For example, Wildlife Computers offers a "black box" version of the HMM with a fixed grid size (0.25 degrees) and minimum speed of movement that may be much faster than the fish is likely to move. This model has so far not been described in the literature, so the details of the likelihood are not known. Another recently developed tool for geolocation of pelagic fishes with the HMM consists of the HMMoce package in R (Braun et al., 2017). This tool has a number of major benefits, such as a thorough description of the data likelihood models in a peer reviewed paper, open source coding, and methods for combining multiple geolocation variables. However, if changes to variance specification methods are desired (e.g., to accommodate the data sources and small-scale environmental heterogeneity of demersal fish applications) they must be implemented by changes to the R code provided.

Much work remains to demonstrate the sensitivity of the model to different choices made for fixed parameters. For example, selection of a method for specifying variance for the distribution of values in each grid cell requires a careful consideration of different approaches. In my experience, the actual variance may be much larger than the variance derived from map heterogeneity, depending on the geolocation variable. In this study, a method to determine variance based on slope values in the study area performed well for the depth variable (Chapter 3), but not the geomagnetic variable (Chapter 4). This is because for some variables, map error is larger than error due to spatial heterogeneity, and this can only be determined by comparing values measured in a known location to mapped values. Future work should

also include a detailed examination of model sensitivity to different values of diffusion, the adequacy of error estimates, and improved ways to assess model performance.

Because it is such a flexible framework, opportunities exist to extend the model in many different directions. For example, if habitat associations are particularly strong for the species studied, integrating a habitat preference map with the data likelihood model could be valuable. Perhaps activity information from accelerometer tags (Nielsen et al., 2018) could be used to augment the movement model with known activity during each time step. Hierarchical models could be very useful for determining diffusion by leveraging values from animals that move through strong environmental gradients, or for which more data are available (some PSAT records have sparse data due to problems with transmission to satellites). In any case, hierarchical approaches are an important step for interpreting movement of individuals in the context of movement of a population (Griffiths et al., 2018).

In addition to methods for characterizing movement, this dissertation has provided observations of Pacific halibut home range behavior (Chapter 1) and migration timing, partial migration, and interannual site fidelity to summer foraging grounds (Chapter 2). The ability to collect this type of detailed information on fish movement over annual timescales has important implications for fisheries management and marine resource planning tools such as time-area closures, identifying the temporal and spatial use of spawning and feeding habitat, and stock delineation/assessment. In addition, it may be helpful for predicting and responding to effects of climate change on the spatial distribution and movement of demersal fish species by tracking changes in habitat use or timing of migrations.

Information on animal space use over time is important for a number of marine spatial planning approaches (Lennox et al., 2018) including restriction of fishing effort at specific times and/or areas. For example, knowledge of the amount of time spent in a closed area is key for Marine Protected Area design and assessing effectiveness (Crossin et al., 2017). In Chapter 2, all tagged adult female Pacific halibut remained within a MPA during the summer and migration out of the MPA occurred when the entire commercial fishery was closed in the winter. Therefore, a combination of spatial closures during summer and temporal closures during winter could provide a measure of protection for Pacific halibut spawning

stock biomass despite their migratory behavior. Electronic tag data has also been used to characterize movement of Atlantic cod and design a Marine Protected Area (Green and Wroblewski, 2000).

Results from the HMM may be aggregated to provide inference about multiple animals. Because each grid cell in the study area contains the probability of occupation by all tagged fish at each time step, and those probabilities can be summarized across specific time periods and for multiple animals, information on the probability of residence for multiple animals in a specific area and time period can be assessed. For example, the HMM has been used to identify spawning areas for Atlantic halibut (Le Bris et al., 2018). If the commercial fishery for Pacific halibut changes to year-round harvests instead of the current winter closures to protect spawning stock, closing certain areas during spawning could be an approach for maintaining protection of Pacific halibut spawning stock. In that case, additional research would be needed to determine whether Pacific halibut from different summer foraging grounds spawn at the same time and place to ensure that winter spawning time-area closures protect Pacific halibut from different summer feeding regions equally.

Stock delineation and stock assessment would also benefit from detailed information on demersal fish movement patterns from methods provided in this dissertation. Information on site fidelity, migration timing, migration extent, and the proportion of the population that participates in migrations can provide insights into potential for spatial structure in populations, which is important in stock assessment activities (Cadrin and Secor, 2009). Observations of site fidelity and homing behavior for Pacific halibut from this dissertation (Chapters 1 and 2) and other studies (Loher, 2008; Seitz et al., 2011), the closely related Atlantic halibut (Le Bris et al., 2018), Pacific cod (Lewandoski et al., 2018), and Atlantic cod (Brander, 1994; Robichaud and Rose, 2001; Robichaud and Rose, 2002; Cote et al., 2004; Tamdrari et al., 2012) suggest that there may be more spatial structure in migratory demersal fish populations than previously thought.

Partial migration is another phenomenon that is important to understand for stock assessment. Partial migration was observed for Pacific halibut, as only 6 out of 21 tagged adult halibut were observed to undertake winter migrations (Chapter 2). The importance of this behavior for stock assessment depends

on whether spawning still occurs if fish do not migrate, or whether residential behavior is an indication that skip-spawning is occurring. None of the detailed data sets (n = 9) for resident fish had evidence of spawning activity (spawning rises) in their depth records, whereas 2 out of 4 detailed data sets from migratory fish showed evidence of spawning activity. If the non-migratory fish were skip-spawners, over-estimation of the spawning biomass could occur if this phenomenon is not accounted for, as predicted for sablefish (Rodgveller et al., 2016). However, it is also possible that resident Pacific halibut spawning behavior differs from migrants and therefore was not detectable in depth records. Future research with other archival tag technology such as high-resolution accelerometer tags may be helpful in identifying spawning behavior that may not be manifested by spawning rises.

Detailed information on movement patterns provided by electronic tags will also be beneficial for understanding potential effects of warming waters due to climate change by providing a mechanistic understanding between movement patterns and temperature. In general, the spatial distribution of fish is expected to shift north as the climate warms (Morley et al., 2018). The spatial distribution of Pacific cod in the Bering Sea has been observed to differ dramatically between cold and warm years (Stevenson and Lauth, 2018). The timing of Atlantic cod spawning advances with warmer temperatures, while evidence of change in spawning locations with temperature is equivocal (McQueen and Marshall, 2017). Understanding whether the Bering Sea Pacific cod population has shifted northward in response to warming temperatures or whether the shifted distribution represents an expansion of the summer feeding area while spawning locations ("home" for cod) remain the same is needed for stock assessment and allocation of quota by area and gear type. For Pacific halibut, for which the summer foraging areas appear to be "home" and migrations are undertaken for spawning, perhaps the timing and destination of spawning migrations would be more variable based on temperature. To understand plasticity in winter spawning vs. summer foraging locations over time or with climate change, tagging durations of more than one year would be needed to determine whether site fidelity occurs for both summer foraging and winter spawning locations.

Finally, the ultimate benefit of detailed movement data provided by electronic tags for fisheries management would be to allow inference at the level of populations. Electronic tags provide detailed information for small numbers of individuals and are extremely important for understanding possible mechanisms that drive movement, spatial distribution, and stock structure. However, extending inference from electronic tagging results to populations given small sample sizes and high variability in behavior among individuals is difficult. Incorporating the HMM geolocation output from a large number of archival-tagged fish into a hierarchical model (Griffiths et al., 2018) may be informative for this purpose, given the hundreds of Pacific halibut archival tag records available for analysis (A.C. Seitz, T. Loher, pers. Comm.). Combining electronic with conventional tags may also be possible, depending on the research question (Le Bris et al., 2013). However, an important future step for translating insights gained from the movement of individuals to populations will be the integration of electronic tag data with methods such as otolith chemistry (Crook et al., 2017) or genetics (Sinclair-Waters et al., 2018) which could allow much larger sample sizes and therefore be more informative at stock and population levels. The methods for obtaining detailed information on the multi-scale movements of demersal fish species described in this dissertation can provide baseline data to design such studies.

References

Brander, K. M. 1994. Patterns of distribution, spawning, and growth in North Atlantic cod: the utility of inter-regional comparisons. ICES marine science symposia, 198: 406-413.

Braun, C. D., Galuardi, B., and Thorrold, S. R. 2017. HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. Methods in Ecology and Evolution, doi: 10.1111/2041-210X.12959.

Cadrin, S. X., and Secor, D. H. 2009. Accounting for spatial population structure in stock assessment: past, present, and future. Pp 405-426 In: B.J. Rothschild and R. Beamish, eds. The Future of Fishery Science in North America. Springer Verlag.

Cote, D., Moulton, S., Frampton, P. C. B., Scruton, D. A., and McKinley, R. S. 2004. Habitat use and early winter movements by juvenile Atlantic cod in a coastal area of Newfoundland. Journal of Fish Biology, 64: 665-679.

Crook, D. A., Buckle, D. J., Allsop, Q., Baldwin, W., Saunders, T. M., Kyne, P. M., Woodhead, J. D., et al. 2017. Use of otolith chemistry and acoustic telemetry to elucidate migratory contingents in barramundi *Lates calcarifer*. Marine and Freshwater Research, 68: 1554-1566.

Crossin, G. T., Heupel, M. R., Holbrook, C. M., Hussey, N. E., Lowerre-Barbieri, S. K., Nguyen, V. M., Raby, G. D., et al. 2017. Acoustic telemetry and fisheries management. Ecological Applications, 27: 1031-1049.

Green, J. M., and Wroblewski, J. S. 2000. Movement patterns of Atlantic cod in Gilbert Bay, Labrador: evidence for bay residency and spawning site fidelity. Journal of the Marine Biological Association of the United Kingdom, 80: 1077-1085.

Griffiths, C. A., Patterson, T. A., Blanchard, J. L., Righton, D. A., Wright, S. R., Pitchford, J. W., and Blackwell, P. G. 2018. Scaling marine fish movement behavior from individuals to populations. Ecology and Evolution, 8: 7031-7043.

Heifetz, J., and Fujioka, J. T. 1991. Movement dynamics of tagged sablefish in the northeastern Pacific. Fisheries Research, 11: 355-374.

Jonsen, I. D., Basson, M., Bestley, S., Bravington, M. V., Patterson, T. A., Pedersen, M. W., Thomson, R., et al. 2013. State-space models for bio-loggers: A methodological road map. Deep Sea Research Part II: Topical Studies in Oceanography, 88–89: 34-46.

Lam, C. H., Nielsen, A., and Sibert, J. R. 2008. Improving light and temperature based geolocation by unscented Kalman filtering. Fisheries Research, 91: 15-25.

Le Bris, A., Fisher, J. A. D., Murphy, H. M., Galbraith, P. S., Castonguay, M., Loher, T., Robert, D., et al. 2018. Migration patterns and putative spawning habitats of Atlantic halibut (*Hippoglossus hippoglossus)* in the Gulf of St. Lawrence revealed by geolocation of pop-up satellite archival tags. ICES Journal of Marine Science, 75: 135-147.

Le Bris, A., Frechet, A., and Wroblewski, J. S. 2013. Supplementing electronic tagging with conventional tagging to redesign fishery closed areas. Fisheries Research, 148: 106-116.

Lennox, R. J., Engler-Palma, C., Kowarski, K., Filous, A., Whitlock, R., Cooke, S. J., and Auger-Méthé, M. 2018. Optimizing marine spatial plans with animal tracking data. Canadian Journal of Fisheries and Aquatic Sciences: 1-13.

Lewandoski, S. A., Bishop, M. A., and McKinzie, M. K. 2018. Evaluating Pacific cod migratory behavior and site fidelity in a fjord environment using acoustic telemetry. Canadian Journal of Fisheries and Aquatic Sciences, 75: 2084-2095.

Liu, C., Cowles, G. W., Zemeckis, D. R., Cadrin, S. X., and Dean, M. J. 2017. Validation of a hidden Markov model for the geolocation of Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences, 74: 1862-1877.

Loher, T. 2008. Homing and summer feeding site fidelity of Pacific halibut (*Hippoglossus stenolepis*) in the Gulf of Alaska, established using satellite-transmitting archival tags. Fisheries Research, 92: 63-69.

Loher, T. 2011. Analysis of match–mismatch between commercial fishing periods and spawning ecology of Pacific halibut (*Hippoglossus stenolepis*), based on winter surveys and behavioural data from electronic archival tags. ICES Journal of Marine Science, 68: 2240-2251.

Loher, T., and Seitz, A. C. 2006. Seasonal migration and environmental conditions of Pacific halibut *Hippoglossus stenolepis*, elucidated from pop-up archival transmitting (PAT) tags. Marine Ecology Progress Series, 317: 259-271.

Maloney, N. E., and Sigler, M. F. 2008. Age-specific movement patterns of sablefish (*Anoplopoma fimbria*) in Alaska. Fishery Bulletin, 106: 305-316.

McQueen, K., and Marshall, C. T. 2017. Shifts in spawning phenology of cod linked to rising sea temperatures. ICES Journal of Marine Science, 74: 1561-1573.

Mitamura, H., Uchida, K., Miyamoto, Y., Kakihara, T., Miyagi, A., Kawabata, Y., Ichikawa, K., et al. 2012. Short-range homing in a site-specific fish: search and directed movements. The Journal of Experimental Biology, 215: 2751.

Morley, J. W., Selden, R. L., Latour, R. J., Frölicher, T. L., Seagraves, R. J., and Pinsky, M. L. 2018. Projecting shifts in thermal habitat for 686 species on the North American continental shelf. PLoS ONE, 13: e0196127.

Musyl, M. K., Brill, R. W., Curran, D. S., Gunn, J. S., Hartog, J. R., Hill, R. D., Welch, D. W., et al. 2001. Ability of archival tags to provide estimates of geographical position based on light intensity. *In* Electronic tagging and tracking in marine fisheries, pp. 343–367. Ed. by J. R. Sibert, and J. L. Nielsen. Kluwer Academic Publs, Dordrecht, The Netherlands.

Nielsen, J. K., Rose, C. S., Loher, T., Drobny, P., Seitz, A. C., Courtney, M. B., and Gauvin, J. 2018. Characterizing activity and assessing bycatch survival of Pacific halibut with accelerometer Pop-up Satellite Archival Tags. Animal Biotelemetry, 6: 10.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. 2008. State-space models of individual animal movement. Trends in Ecology & Evolution, 23: 87-94.

Pedersen, M. W., Righton, D., Thygesen, U. H., Andersen, K. H., and Madsen, H. 2008. Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. Canadian Journal of Fisheries and Aquatic Sciences, 65: 2367-2377.

Robichaud, D., and Rose, G. A. 2001. Multiyear homing of Atlantic cod to a spawning ground. Canadian Journal of Fisheries and Aquatic Sciences, 58: 2325-2329.

Robichaud, D., and Rose, G. A. 2002. The return of cod transplanted from a spawning ground in southern Newfoundland. ICES Journal of Marine Science, 59: 1285-1293.

Rodgveller, C. J., Stark, J. W., Echave, K. B., and Hulson, P.-J. F. 2016. Age at maturity, skipped spawning, and fecundity of female sablefish (*Anoplopoma fimbria)* during the spawning season. Fisheries Bulletin, 114: 89-102.

Schaefer, K. M., Fuller, D. W., and Block, B. A. 2011. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the Pacific Ocean off Baja California, Mexico, determined from archival tag data analyses, including unscented Kalman filtering. Fisheries Research, 112: 22-37.

Seitz, A. C., Loher, T., Norcross, B. L., and Nielsen, J. L. 2011. Dispersal and behavior of Pacific halibut *Hippoglossus stenolepis* in the Bering Sea and Aleutian Islands region. Aquatic Biology, 12: 225-239.

Seitz, A. C., Norcross, B. L., Wilson, D., and Nielsen, J. L. 2006. An evaluation of light-based geolocation for demersal fish in high latitudes. Fishery Bulletin, 104: 571-578.

Shimada, A., and Kimura, D. 1994. Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. Fishery Bulletin, 92: 800-816.

Sinclair-Waters, M., Bentzen, P., Morris, C. J., Ruzzante, D. E., Kent, M. P., Lien, S., and Bradbury, I. R. 2018. Genomic tools for management and conservation of Atlantic cod in a coastal marine protected area. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1915-1925.

Skud, B. E. 1977. Drift, migration, and intermingling of Pacific halibut stocks. Scientific Report No. 63, International Pacific Halibut Commission, Seattle, WA. 42 pp.

Stevenson, D. E., and Lauth, R. R. 2018. Bottom trawl surveys in the northern Bering Sea indicate recent shifts in the distribution of marine species. Polar Biology.

Strøm, J. F., Thorstad, E. B., Chafe, G., Sørbye, S. H., Righton, D., Rikardsen, A. H., and Carr, J. 2017. Ocean migration of pop-up satellite archival tagged Atlantic salmon from the Miramichi River in Canada. ICES Journal of Marine Science, 74: 1356-1370.

Tamdrari, H., Brêthes, J. C., Castonguay, M., and Duplisea, D. E. 2012. Homing and group cohesion in Atlantic cod *Gadus morhua* revealed by tagging experiments. Journal of Fish Biology, 81: 714-727.

Thompson, G. G., and Dorn, M. W. 2004. Assessment of the Pacific cod stock in the eastern Bering Sea and Aleutian Islands area. Pp 185-302 In: Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Regions. Anchorage, AK: North Pacific Fishery Management Council.

Thygesen, U., Pedersen, M., and Madsen, H. 2009. Geolocating Fish Using Hidden Markov Models and Data Storage Tags. *In* Tagging and Tracking of Marine Animals with Electronic Devices, pp. 277-293. Ed. by J. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert. Springer Netherlands.

Appendix A: Geolocation of demersal fishes in the North Pacific Ocean: Hidden Markov model framework and data likelihood models[1]

A1. Abstract

Understanding large-scale movements of demersal fishes is important for fisheries management, but detailed information on large-scale movement patterns is scarce due to the difficulty of obtaining it with methods that provide only release and recovery locations. However, archived data from electronic tags can be used to estimate locations of tagged fish at each time step through the use of geolocation models that link data measured by the tag to maps of the study area. State-space models that use archival tag data such as depth, temperature, and light to reconstruct movement pathways are a powerful tool for obtaining information on migration pathways or spawning locations. We adapted a hidden Markov model (HMM) developed for Atlantic cod (*Gadus morhua*) in the North Sea for geolocation of demersal fishes in the North Pacific Ocean. The data likelihood model is based on the maximum depth in each time bin, but other geolocation data can be utilized when available. Here, we provide a detailed description of the data likelihood model, HMM framework, and resulting model products. We intend this document to serve as a reference for use of the HMM for geolocation of demersal fish species in the North Pacific Ocean as well as provide an example of adapting the HMM for other applications.

A2. Introduction

A hidden Markov model (HMM) is a type of state-space model (SSM) commonly used to model time series processes with discrete states. For example, a movement process for humans could include the states of standing, walking, or running. In an HMM, states cannot be observed directly but indirect aspects of the process can be observed (e.g., velocity in the human movement example). Indirect observations are assumed to be linked to the hidden state in a probabilistic way. For example, if the average human running speed is 15 km/h and walking is 5 km/h, information on the velocity of an unobserved human at any given time can be used to infer whether the person is standing, walking, or running.

[1] Nielsen, J. K., Mueter, F., Adkison, M., Loher, T., McDermott, S., and Seitz, A. C.

In general, an HMM predicts the future state of a process based on an estimation of the current state, knowledge of transition probabilities to a different state, and indirect observations of the hidden process at each time step (Zucchini and MacDonald, 2009). As applied to fish geolocation models, the location of the tagged animal at discrete time intervals between tag release and tag recovery is the unobserved "state" estimated by the model. Transition probabilities are determined by a movement model that reflects the hypothesized underlying movement pattern of the animal in the study area. Data collected by the archival tag provide indirect observations of the hidden state (location) of the tagged fish when they are matched to maps of geolocation variables in the study area. Thus, the HMM has the same form as other SSMs in that it consists of an underlying movement (process) model coupled with a data likelihood (observation) model (Patterson et al., 2008); the difference between the HMM and other SSMs is that the HMM consists of discrete vs. continuous states.

The HMM developed for demersal fish species (Pedersen, 2007; Pedersen et al., 2008; Thygesen et al., 2009) provides a general methodological framework that can be adapted for application to other fish species as well. Matlab code for the model was provided by Martin Pedersen (DTU Aqua, Denmark). This code serves as the basis of the current application of the HMM to demersal fishes as described in this appendix.

A3. Model description

A3.1 Overview

The key feature of the HMM is the partition of the study area into discrete grid cells. This grid is represented in the model by a matrix $\mathbf{\Phi}$ with i rows and j columns (corresponding to latitude and longitude); each element i,j contains the probability that the fish is located in that grid cell at a given time step k. One of the major products of the HMM is an expansion of this study area matrix to a 3-dimensional array that contains the probability distribution of the tagged fish in each grid cell of the study area at each time step. This 3-D array $\mathbf{\Phi}_{ijk}$ is produced by 1) a recursive forward filter that begins at the time that the tagged fish is released, matches the predicted locations of the fish from the movement model to data collected by the tagged fish at each time step, and ends at the time that the tag is recovered, and 2)

a backward smoothing recursion that begins at the tag recovery location and "re-weights" the probabilities from the forward filter at each time step based on knowledge of the recovery location. Because the position estimates are derived from probability values within each grid cell, the resolution of position estimates depends on the grid size chosen for the model.

A3.2 Data likelihood model for demersal fishes

The data likelihood model is a core component of a geolocation model. At any given location, it provides the probability of observing the environmental data recorded by a fish (e.g., depth or temperature) given the model used to describe the distribution of values at that location (e.g., mean and variance from bathymetric or bottom temperature maps). Here, we describe a data likelihood model for demersal fishes that is focused on depth data, but can include other geolocation variables when available. We discuss different options for depth-based likelihoods and the ways in which depth likelihoods can be combined with other geolocation variables (e.g., light-based longitude and known locations from acoustic telemetry).

A3.2.1 Depth-based likelihoods

Depth data are provided by most electronic archival tags, although data format and quality (accuracy/precision) may differ by tag type, manufacturer, and depth range in the deployment area. Depth may be obtained from archival tags that are physically recovered as a high-resolution time series, with measurement intervals of seconds to minutes (referred to subsequently as "detailed data"). Depth may also be available as summarized data transmitted by Pop-up Satellite Archival Tags (referred to subsequently as "PSAT data"). For example, PSAT data may consist of minimum, maximum, or average depths recorded over 12- or 24-hour time periods. Electronic tags deployed in deeper water may have less precision than tags deployed in shallow water, as resolution is typically specified as a percent of the tag's depth range.

Likelihoods based on depth vary depending on the type of tag data available (detailed vs. PSAT), the magnitude and orientation of environmental gradients in the study area, and fish behavior. For example, the original HMM developed for Atlantic cod featured two types of likelihoods based on

detailed depth data from physically-recovered archival tags (Pedersen et al., 2008). On days when tagged fish were stationary, the rise and fall of the tide was recorded in tag depth records. Potential locations of the tagged fish could then be determined by matching the amplitude and phase of the tide signal recorded by the fish to amplitude and phase predicted in different parts of the study area by oceanographic models (Hunter et al., 2004). This is known as a "tidal likelihood." On days when the fish was more active in the water column, and thus no tidal signal was available, the likelihood was obtained by matching the maximum depth per time period to maps of bathymetry in the study area. Because demersal fishes are likely to contact the seafloor at least once per 12- or 24-hour tag-recording period, the maximum depth recorded by a tagged fish during that time period can be assumed to represent the depth of the seafloor in the inhabited location. This is known as a "maximum depth likelihood."

The data likelihood model described here for geolocation of demersal fishes in the North Pacific Ocean is based on maximum depth likelihood. Bathymetry gradients can be steep in Alaska, and bathymetric map resolution and accuracy are improving steadily in many areas (Zimmermann and Prescott, 2015). In contrast, the tidal likelihood is less useful in Alaska as gradients in tide amplitude and phase tend to be weaker than the North Sea and models of tide amplitude and phase are not well-developed. In addition, coarse depth resolution of tags deployed in deep waters may make detection of tidal patterns in the data more difficult. However, in certain areas such as southeastern Alaska (Shi et al., 2014), tidal models may be sufficiently accurate to allow inclusion of tidal likelihood in the data likelihood model. Because the maximum depth likelihood can accommodate both detailed and PSAT depth data sets, a data likelihood based on maximum depth likelihood can be used for most demersal fish data sets in Alaska.

For the maximum depth likelihood, tag data can be matched to bathymetry measurements in a number of different ways. Specific likelihood models differ based on the behavior of the fish (whether it is assumed to be directly on the bottom, anywhere in the water column, or within some distance of the bottom), the statistical distribution of depth values in the study area (e.g., normal or log-normal), and

whether or not uncertainty in tag measurement is included in the model. Following is a discussion of three maximum depth likelihood methods (Figure A1) that could be used for demersal fish.

A3.2.1.1 Normal Cumulative Density Function likelihood

The maximum depth likelihood employed for the original HMM (Pedersen et al., 2008) is obtained from a normal cumulative distribution function (CDF) of the mean depth and estimated depth variance within each grid cell. The likelihood value is the CDF quantile represented by the tag depth, normalized by a CDF truncated at a depth of zero:

$$L = \Phi\left[\frac{z_{max} - z(x)}{\sigma(x)}\right]\left(1 - \Phi\left[\frac{-z(x)}{\sigma(x)}\right]\right)^{-1} \quad \text{(eq A1)}$$

where Φ is a Gaussian cumulative distribution function, z(x) is the depth of the grid cell, z_{max} is the maximum observed depth per time step, and σ(x) is the standard deviation of the bathymetry. This likelihood assumes that the fish could be anywhere in the water column when the maximum depth was recorded. Probability increases for depths greater than the tag depth measurements and decreases for depths shallower than the tag measurement as a Gaussian function of bathymetry uncertainty (Figure A1-A). This likelihood method may not be optimal for species such as flatfishes, which are frequently in direct contact the sea floor (Gibson, 2005) and therefore probability on both sides of the mean cell depth would decline as a function of uncertainty in the bathymetry. In addition, there is no way to incorporate tag measurement resolution into the likelihood value.

A3.2.1.2 Integrated normal Probability Density Function likelihood

A maximum depth likelihood specification introduced in an application of the HMM to the geolocation of Atlantic cod in the Gulf of St. Lawrence (Le Bris et al., 2013) addresses both of these issues. This likelihood is calculated by integrating a normal probability density function (PDF) of the mean depth and estimated depth variance within each grid cell between the interval of the measured depth ± tag measurement resolution:

$$L = \int_{z_1}^{z_2} N(z; \mu_z, \sigma_z)\, dz \quad \text{(eq. A2)}$$

where z is maximum depth measured by the fish during the time interval, z_1 and z_2 are the lower and upper limits of uncertainty in tag measurement, μ_z is the mean of the bathymetry grid cell, and σ_z is the standard deviation of depths within the grid cell (Figure A1-B).

A3.2.1.3 Integrated split normal Probability Density Function likelihood

A likelihood based on a split normal distribution would be appropriate for demersal fishes located close to the sea floor, but not necessarily in contact with it (Webber, 2015). For example, an archival tagging study of Pacific cod indicated that most tagged fish were likely within 10 m of the ocean floor during the day (Nichol et al., 2007). A split normal distribution allows the specification of separate standard deviations for each half of the PDF:

$$f(x) = \begin{cases} A\, exp\left[\frac{-(x-\mu)^2}{2\sigma_1^2}\right] & x \leq \mu \\ A\, exp\left[\frac{-(x-\mu)^2}{2\sigma_2^2}\right] & x \geq \mu \end{cases} \quad \text{(eq A3)}$$

where $A = (\sqrt{2\pi}(\sigma_1 + \sigma_2)/2)^{-1}$, μ is the bathymetry cell mean, σ_1 is the standard deviation for the left (shallow) side of the bathymetry PDF, and σ_2 is the standard deviation for the right (deeper) side of the bathymetry PDF (Wallis, 2014). Therefore, the probability distribution accounts for bathymetric uncertainty on the deeper half of the grid cell mean; and both bathymetric uncertainty and the probability that the fish could be slightly above the seafloor on the shallower half (Figure A1-C). The likelihood value can be obtained by integration between the limits of tag measurement resolution in the same manner as the integrated normal PDF likelihood.

A3.2.2 Other geolocation variables

Although the data likelihood model for demersal fishes in Alaska is based on maximum depth, additional geolocation variables may be available that could be combined to increase precision and accuracy of geolocation models. For example, light-based longitude may be available for demersal fish

that occupy depths < 150 m (Seitz et al., 2006). If maps of bottom temperature or accurate maps of tidal amplitude/phase become available, these can also be included in the likelihood model. In certain areas, newly-developed tags that measure the Earth's magnetic field (Klimley et al., 2017) could also provide a geomagnetic likelihood. The ability to combine multiple geolocation variables will improve geolocation estimates, particularly when gradients for different geolocation variables are orthogonal.

A3.2.3 Known locations or regions

The HMM can easily incorporate known locations provided by acoustic telemetry or other methods into the data likelihood model by assigning positive values to grid cells within known locations and zero values to grid cells outside known locations. For example, for a fish double-tagged with acoustic tags (or SPOT tags, if fish come to the surface), grid cells within the range of the location estimate plus telemetry error can be assigned a value of 1 while all other cells in the study area are assigned a value of zero. Continuous probability values within known regions can also be assigned to grid cells, e.g. based on distance from fixed telemetry receivers (Pedersen and Weng, 2013).

A3.2.4 Combining likelihoods

Likelihoods for different variables can be combined by cell-wise multiplication, as long as they are not correlated (Le Bris et al., 2013; Liu et al., 2017).

$$L_{total} = L_{depth} * L_{longitude} * L_{telemetry} \quad \text{(eq A4)}$$

If all data for a given type of likelihood are missing in a given time step, all grid cells for that time step are assigned a value of one.

A4. Forward filter

The model is initiated by assigning a value of one to the grid cell where the tagged fish was released ($\mathbf{\Phi}_{k=1}$) and a value of zero to all other cells on the grid. To obtain the probability distribution for $\mathbf{\Phi}_{k=2}$, $\mathbf{\Phi}_{k=1}$ is updated first by the movement model, which consists of convolution with the diffusion kernel H, and then by the data likelihood model for time k=2 (Figure A2). The posterior for time k=2 then

becomes the prior for time k=3, and the model proceeds with alternating updates from the movement and data likelihood model until the final time step is reached. The formula for this recursion is

$$\mathbf{\Phi}_{k+1} = \mathbf{\Phi}_{k} * \mathbf{H}_{k:k+1} \times \mathbf{L}_{k+1} \quad \text{(eq A5)}$$

where $\mathbf{\Phi}_k$ is the matrix that contains the estimated probability distribution in the study area at time k, $\mathbf{H}$ is a matrix that contains the transition probabilities from the state at time k to the state at time k+1, $\mathbf{L}_{k+1}$ is a matrix containing the data likelihood at time k+1, * denotes matrix convolution, and x denotes element-wise multiplication.

The underlying movement model is assumed to be isotropic diffusion, for which the forward Kolmogorov equation (also known as Fokker-Planck equation) is used to distribute the movement probability to each cell at a future time step by finite-differences (Pedersen et al., 2008). As implemented in the Matlab code, this is achieved by convolution of the matrix $\mathbf{\Phi}$ at time k with a two-dimensional Gaussian diffusion kernel ($\mathbf{H}$) to obtain the predicted distribution of the fish at the next time step (k+1). The size of the diffusion kernel varies based on the value of diffusion (D). Different movement states (such as limited movement during foraging vs. greater movement during migration) may be assigned to each time step based on auxiliary analyses prior to modeling. D may be directly specified based on knowledge of fish movement rates from previous research, or it may be estimated by maximum likelihood in the HMM (see below).

After the update from the movement model, the probability distribution is updated by the data likelihood model. The data likelihood represents the probability of observing the data recorded by the tag at time k given the values (e.g., mean and s.d.) present within each cell of the study area, as described above for demersal fish. Following the data likelihood update, probability distribution is normalized by dividing each grid cell by the sum of all grid cells, denoted by λ_k. The normalized probabilities serve as the prior for the next time step in the recursion.

A5. Estimating diffusion

The normalization constant λ_k reflects how well the data likelihood fits the predicted locations at each time step, and therefore it is used as the basis of the maximum likelihood estimation of the size of the diffusion kernel. The forward recursion is repeated with different values of diffusion until the negative log likelihood value is minimized:

$$L(\theta) = \sum_{k=1}^{n} -\log(\lambda_k) \quad \text{(eq A6)}$$

where θ is the value of diffusion (D) used for the movement model and λ_k is the normalization constant obtained at each time step k of the forward recursion. The negative log likelihood is minimized using the fminbnd command in Matlab 2016. The uncertainty in the estimate of D is also provided by the Hessian matrix. The negative log likelihood can also be used to assess performance with one movement state versus two using AIC.

A6. Backward smoothing

After the forward recursion is completed using the maximum likelihood estimate of D (or assigned value of D from auxiliary sources), backward smoothing is initiated. Backward smoothing consists of a backward recursion that begins at the tag recovery location and adjusts the probabilities obtained from the forward recursion by incorporating knowledge of the end location:

When k = time of tag release or time of tag recovery,

$\text{Smooth}_k = \mathbf{\Phi}_k$

For all other k,

$$\text{Smooth}_k = \mathbf{\Phi}_k \times (\text{Smooth}_{k+1}/\text{Predicted}_{k+1} * \mathbf{H}) \quad \text{(eq A7)}$$

where $\mathbf{\Phi}_k$ is the matrix that contains the normalized estimated probability distribution at time k resulting from the forward filter, **H** is the diffusion kernel that contains the transition probabilities from the state at time $k+1$ to the state at time k, Predicted_{k+1} refers to the probability matrix after the movement update in

the forward filter at time k (e.g., $\mathbf{\Phi}_k$ * $\mathbf{H}_{k:k+1}$ in eqn. A1), * denotes matrix convolution, and × denotes element-wise multiplication. The smoothed 3-D array $\mathbf{\Phi}_{ijk}$ that results from this process contains the posterior distribution for each grid cell at each time step given all of the information available in the data set, or $P(X_k|Y_{k=1:n})$. This probability array can then be used to calculate residency distributions over given periods of time or simulate random tracks to obtain the probability of residence in specific areas.

A7. Movement states

Different rates of movement during different movement states, such as foraging versus migrating, can be accommodated in the model by specifying different sizes of the diffusion kernel. To preserve the generality and simplify the model, movement states are specified prior to estimation as either a low or a high rate of movement. In some applications, the presence of a tidal pattern in depth data is an indication of relative rate of movement (Pedersen et al., 2008; Liu et al., 2017). Movement states can also be determined by auxiliary data such as acceleration (Nielsen et al., 2018). The fit of the model with different movement state treatments can be assessed using AIC based on the negative log likelihood calculated in Equation A6.

A8. Model products

A8.1 Smoothed estimates

The 3-D array of smoothed estimates $\mathbf{\Phi}_{ijk}$ is the primary product of the model. This array provides the probability that the fish is located in each grid of the study area at each time step given all available data. To display this information, either probability in the study area at specific time steps can be shown, or movies that show the progression of probability over time from the release date to day of recovery can be produced.

A8.2 Residency distribution

The smoothed estimates $\mathbf{\Phi}_{ijk}$ can be used to calculate residency distributions that describe the estimated residence time within each grid cell over a specific time interval (Pedersen et al., 2011). Residency distributions are obtained by summing probabilities for each grid cell at each regularly-spaced time step within the time period of interest (e.g., for the entire archival record, by month, by movement

state, or for day vs. night) and plotting by quantile distribution (Figure 3A). For example, a residency distribution could identify the area where a tagged fish was likely to have spent 50% of its time over the course of the study.

A8.3 Reconstructing movement pathways

Although the smoothed estimates $\mathbf{\Phi}_{ijk}$ do provide information on the spatial distribution of possible locations for the tagged fish at each time step, an estimated sequence of locations (e.g., a reconstructed movement pathway) is also an important objective for geolocation models. Movement pathways (Figure 3B) can be reconstructed by global decoding with the Viterbi algorithm or by local decoding based on the smoothed estimates at each time step. These approaches may differ in terms of accuracy, precision, and processing time depending on HMM grid size and environmental gradient strength and degree of habitat heterogeneity (Chapter 3).

A8.3.1 Most Probable Track

One option for reconstructing movement paths consists of using the Viterbi algorithm (Viterbi, 1967) to determine the sequence of states that has the highest probability in an HMM (Zucchini and MacDonald, 2009); this sequence of states is referred to as the most probable track (MPT). It can be thought of as global decoding of the posterior distribution (Pedersen et al., 2011) that is based on the formula for calculating the overall probability of observing a particular sequence of states:

$$P(X_{k=1:n},Y_{k=1:n}) = P(X_{k=1})\, P(Y_{k=1}|X_{k=1}) \prod_{k=2}^{n} P(X_k|X_{k-1})\, P(Y_k|X_k) \qquad \text{(eq A8)}$$

where $P(X_{t=1:n},Y_{t=1:n})$ is the joint probability of the hidden states (X) and the indirect observations of the states (Y), n is the total number of observations, $P(X_{k=1})$ is the probability of the initial state, $P(Y_{k=1}|X_{k=1})$ is the probability of the first observation, $P(X_k|X_{k-1})$ is the transition probability from X_{k-1} to X_k, and $P(Y_k|X_k)$ is the data likelihood at time step k. The Viterbi algorithm is a recursion that calculates the sequence of states $X_{1:k-1}$ that provides the greatest probability for each possible state (e.g., grid cell) at time step k. Starting at the release location, the transition probabilities are first applied (convolution with

the diffusion kernel) followed by element-wise multiplication by the data likelihood in the same manner as the forward filter. However, instead of normalizing the probability at the end of each time step, the algorithm begins the recursion at time step k+1 using the magnitude of the highest sequence probability for each state (e.g., grid cell) at time step k and also records which previous state (X_{k-1}) provided the highest probability for each state at time step k. At the end of the recursion, the sequence with the highest probability will be represented by the state (grid cell) with the highest probability at the last time step. The sequence of states with the highest probability can then be recovered based on the recorded information on which state provided the highest probability at each previous time step.

The Viterbi algorithm has been used to reconstruct movement paths with the HMM in several fish geolocation studies (Pedersen et al., 2008; Thygesen et al., 2009; Woillez et al., 2016; Liu et al., 2017; Strøm et al., 2017). However, this method has several disadvantages. First, the calculation can be time consuming if the number of grid cells is large. Second, the MPT may diverge from the smoothed probability estimates in certain cases, such as when environmental variable gradients are weak and the underlying movement model has a greater influence than the data likelihood model (Pedersen et al., 2011). Third, it is difficult to quantify uncertainty in the MPT.

A8.3.2 Weighted mean probability

As an alternative to the global decoding provided by the Viterbi algorithm, some researchers have used local decoding methods based on the maximum, mean, or mode of the smoothed estimate distribution $\mathbf{\Phi}_{ijk}$ at each time step to reconstruct movement paths (Thygesen et al., 2009; Woillez et al., 2016; Biais et al., 2017; Braun et al., 2017). In an approach employed in two recent studies (Biais et al., 2017; Doherty et al., 2017), movement paths are reconstructed by calculating the mean latitude and longitude from all grid cells in the study area weighted by the smoothed probability distribution at each time step.

A8.4 Probability quantile polygons

Uncertainty in the weighted mean movement paths is also derived from the smoothed probability distribution at each time step. Previous studies (Biais et al., 2017; Braun et al., 2017) have quantified

uncertainty by fitting a 95% error ellipse to the weighted variance/covariance at each time step. However, the resulting error ellipses do not always fully reflect the contours and gridded nature of the smoothed probability estimates. We propose a new method for visualizing and quantifying uncertainty by constructing polygons that incorporate specified quantiles of the gridded probability surface (Figures A4 and A5).

A8.5 Diffusion estimates

The model is capable of calculating the optimal size of the diffusion kernel that serves as the movement model. However, the accuracy of this estimate is related to model grid size, study area gradients, and degree of study area heterogeneity (Chapter 3). In low-gradient study areas or in cases with substantial missing data, using diffusion rates based on prior knowledge should be considered if such data are available. If estimated in the model, the value of diffusion may not be equated to animal movement speed, as fish rarely have purely diffusive movements. Obtaining step-length distributions from reconstructed movement pathways may provide a more realistic description of fish movement speed based on model estimates. Examination of step-length distributions is also a way to assess whether or not the model results are realistic.

A9. Conclusion

The HMM developed by Pedersen et al. (2008) has proven to be a remarkably flexible tool for geolocation of many types of fish in different geographic regions. The model is adapted by choosing the optimal data likelihood model, selecting the appropriate grid size, and deciding whether to estimate the size of the diffusion kernel (e.g., the movement model) with the model or use a specified value based on auxiliary data. These decisions all require careful consideration of underlying assumptions (e.g., a normal distribution of depths in each bathymetry grid cell) relative to degree of heterogeneity in geolocation variables (Chapter 3). The model framework and data likelihood model described here can easily be modified for application to pelagic species as well. Additional model details for pelagic applications are also available in Braun et al. (2017).

A10. Literature cited

Biais, G., Coupeau, Y., Séret, B., Calmettes, B., Lopez, R., Hetherington, S., and Righton, D. 2017. Return migration patterns of porbeagle shark (*Lamna nasus*) in the Northeast Atlantic: implications for stock range and structure. ICES Journal of Marine Science, 74: 1268-1276.

Braun, C. D., Galuardi, B., and Thorrold, S. R. 2017. HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. Methods in Ecology and Evolution, doi: 10.1111/2041-210X.12959.

Doherty, P. D., Baxter, J. M., Gell, F. R., Godley, B. J., Graham, R. T., Hall, G., Hall, J., et al. 2017. Long-term satellite tracking reveals variable seasonal migration strategies of basking sharks in the north-east Atlantic. Scientific Reports, 7: 42837. doi: 42810.41038/srep42837.

Gibson, R. N. 2005. The behaviour of flatfishes. *In* Flatfishes: biology and exploitation, pp. 213-239. Ed. by R. N. Gibson. Blackwell Science Ltd, Oxford, UK.

Hunter, E., Metcalfe, J. D., Holford, B. H., and Arnold, G. P. 2004. Geolocation of free-ranging fish on the European continental shelf as determined from environmental variables II. Reconstruction of plaice ground tracks. Marine Biology, 144: 787-798.

Klimley, A. P., Flagg, M., Hammerschlag, N., and Hearn, A. 2017. The value of using measurements of geomagnetic field in addition to irradiance and sea surface temperature to estimate geolocations of tagged aquatic animals. Animal Biotelemetry, 5: 19.

Le Bris, A., Frechet, A., and Wroblewski, J. S. 2013. Supplementing electronic tagging with conventional tagging to redesign fishery closed areas. Fisheries Research, 148: 106-116.

Liu, C., Cowles, G. W., Zemeckis, D. R., Cadrin, S. X., and Dean, M. J. 2017. Validation of a hidden Markov model for the geolocation of Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences, 74: 1862-1877.

Nichol, D., Honkalehto, T., and Thompson, G. 2007. Proximity of Pacific cod to the sea floor: Using archival tags to estimate fish availability to research bottom trawls. Fisheries Research, 86: 129-135.

Nielsen, J. K., Rose, C. S., Loher, T., Drobny, P., Seitz, A. C., Courtney, M. B., and Gauvin, J., 2018. Characterizing activity and assessing bycatch survival of Pacific halibut with accelerometer Pop-up Satellite Archival Tags. Animal Biotelemetry 6:10. DOI 10.1186/s40317-018-0154-2.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. 2008. State-space models of individual animal movement. Trends in Ecology & Evolution, 23: 87-94.

Pedersen, M. W. 2007. Hidden Markov models for geolocation of fish. Technical University of Denmark, Kongens Lyngby, Denmark.

Pedersen, M. W., Patterson, T. A., Thygesen, U. H., and Madsen, H. 2011. Estimating animal behavior and residency from movement data. Oikos, 120: 1281-1290.

Pedersen, M. W., Righton, D., Thygesen, U. H., Andersen, K. H., and Madsen, H. 2008. Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. Canadian Journal of Fisheries and Aquatic Sciences, 65: 2367-2377.

Pedersen, M. W., and Weng, K. C. 2013. Estimating individual animal movement from observation networks. Methods in Ecology and Evolution, 4: 920-929.

Seitz, A. C., Norcross, B. L., Wilson, D., and Nielsen, J. L. 2006. An evaluation of light-based geolocation for demersal fish in high latitudes. Fishery Bulletin, 104: 571-578.

Shi, L., Wang, J., Myers, E., and Huang, L. 2014. Development and use of tide models in Alaska supporting VDatum and hydrographic surveying. Journal of Marine Science and Engineering, 2: 171-193.

Strøm, J. F., Thorstad, E. B., Chafe, G., Sørbye, S. H., Righton, D., Rikardsen, A. H., and Carr, J. 2017. Ocean migration of pop-up satellite archival tagged Atlantic salmon from the Miramichi River in Canada. ICES Journal of Marine Science, 74: 1356-1370.

Thygesen, U., Pedersen, M., and Madsen, H. 2009. Geolocating Fish Using Hidden Markov Models and Data Storage Tags. *In* Tagging and Tracking of Marine Animals with Electronic Devices, pp. 277-293. Ed. by J. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert. Springer Netherlands.

Viterbi, A. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory, 13: 260-269.

Wallis, K. F. 2014. The two-piece normal, binormal, or double Gaussian distribution: its origin and rediscoveries. Statistical Science, 29: 106-112.

Webber, D. N. 2015. Modelling complexity and uncertainty in fisheries stock assessment. Victoria University of Wellington.

Woillez, M., Fablet, R., Ngo, T.-T., Lalire, M., Lazure, P., and de Pontual, H. 2016. A HMM-based model to geolocate pelagic fish from high-resolution individual temperature and depth histories: European sea bass as a case study. Ecological Modelling, 321: 10-22.

Zimmermann, M., and Prescott, M. M. 2015. Smooth Sheet Bathymetry of the Central Gulf of Alaska.

Zucchini, W., and MacDonald, I. 2009. Hidden Markov models for time series: an introduction using R, Chapman & Hall/CRC, Boca Raton, FL.

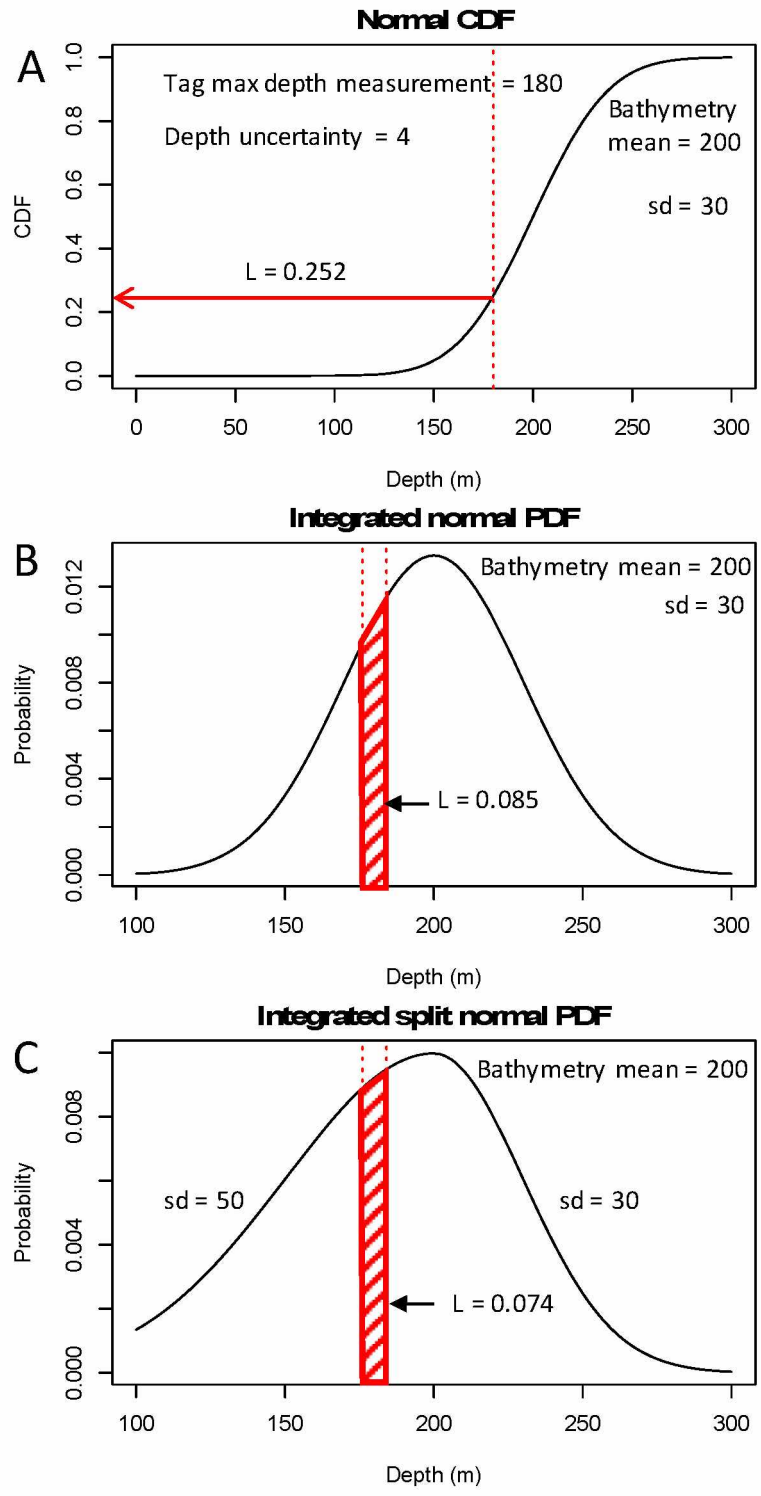


Figure A1. Options for maximum depth data likelihood models that link tag data to study area bathymetry grid cells. A) Normal CDF method (Pedersen et al., 2008) eliminates shallow depths when fish could be anywhere in the water column. B) Integrated normal PDF (Le Bris et al., 2013) links tag depth to bathymetry when fish is assumed to be on the seafloor. C) Integrated split normal PDF (Webber, 2015) allows some probability for shallower depths when fish is likely to be near the seafloor but not in direct contact.

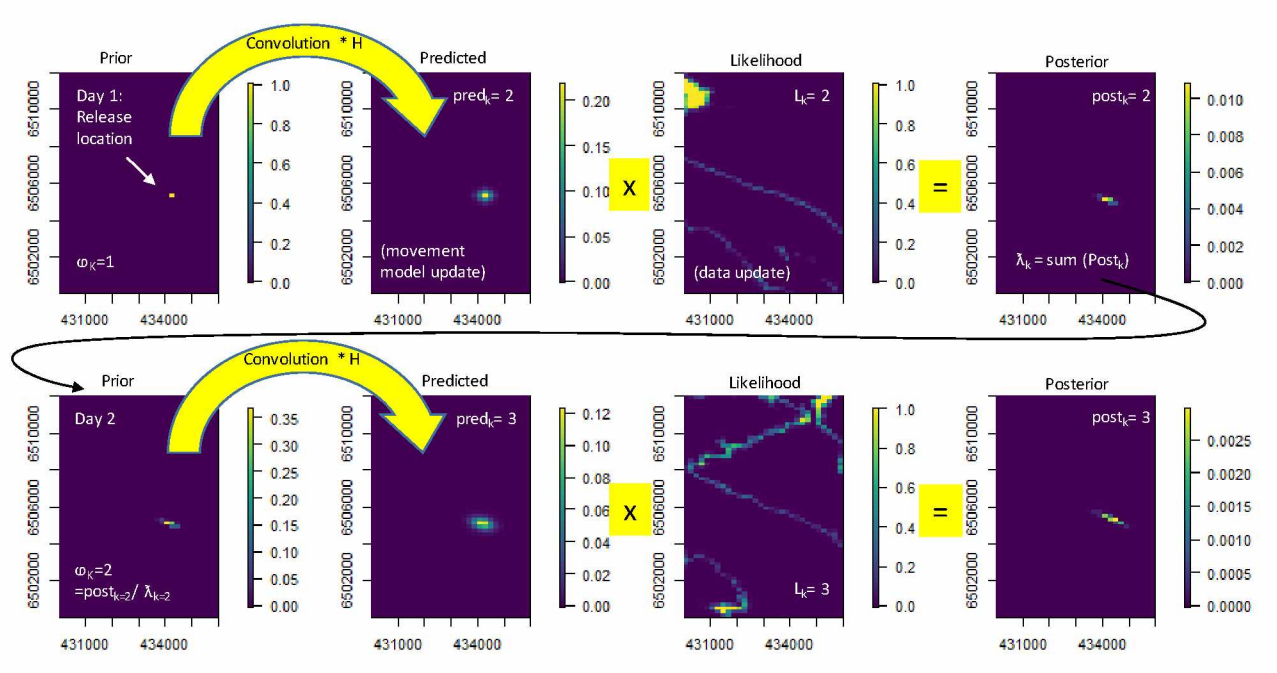


Figure A2. Example of the forward filter for the first two time steps. The model starts with all of the probability in the grid cell where the tagged fish was released on Day 1 ($\Phi_{k=1}$), which is the prior distribution for Day 2. A movement update of the prior is accomplished by convolution with an isotropic diffusion kernel (H) to obtain the fish's predicted location on Day 2. A data update is accomplished by element-wise multiplication of the predicted ("pred") and data likelihood ("L") matrices for Day 2. This update results in the posterior for Day 2, which is normalized by lambda, the sum of the posterior, and becomes the prior for Day 3. Lambda is the quantity used to determine log likelihood (Eq. A6; see text). The model continues in this manner until the day that the tag is recovered.

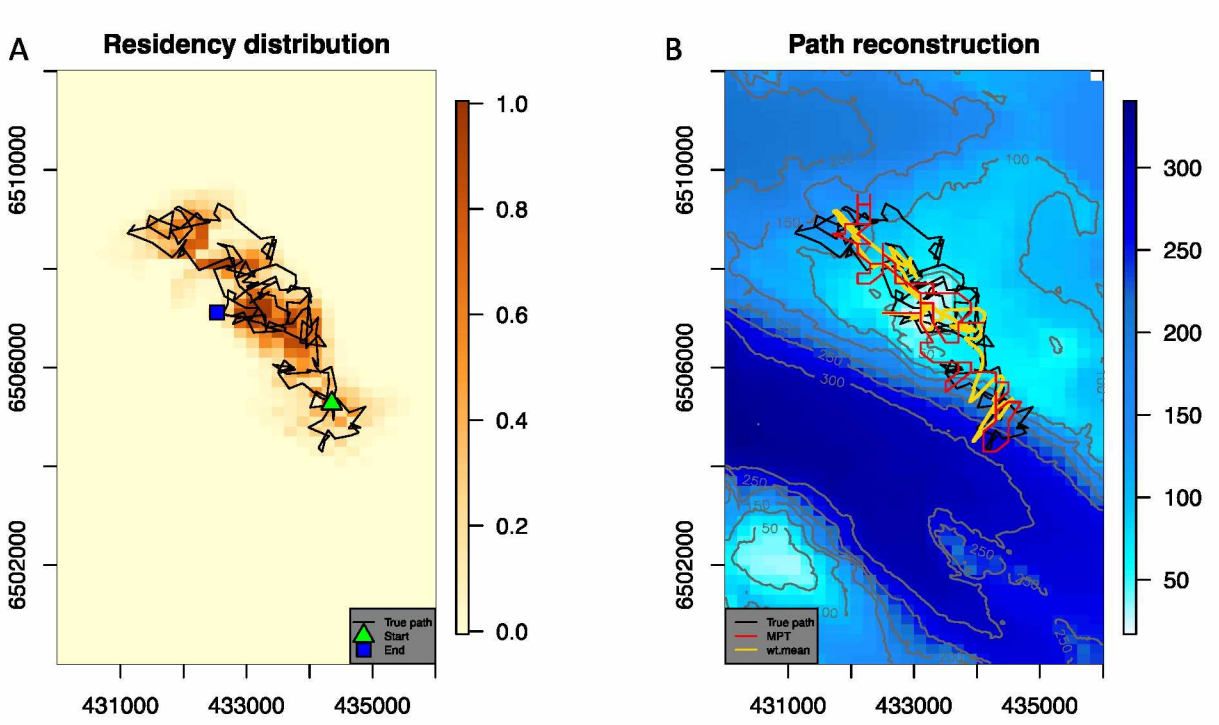


Figure A3. Example of A) residency distribution and B) movement path reconstruction for a simulated path with 200 steps (black line). A) Residency distribution (color-coded surface) is the sum of probabilities in each grid cell over all 200 steps and color-coded by probability quantiles (highest cumulative probability in darker colors). The initial location is a green triangle, and the final location is a blue square. B) Two methods of reconstructing movement paths are shown on a 200 m bathymetry grid: most probable track (red line) and weighted mean probability (yellow line).

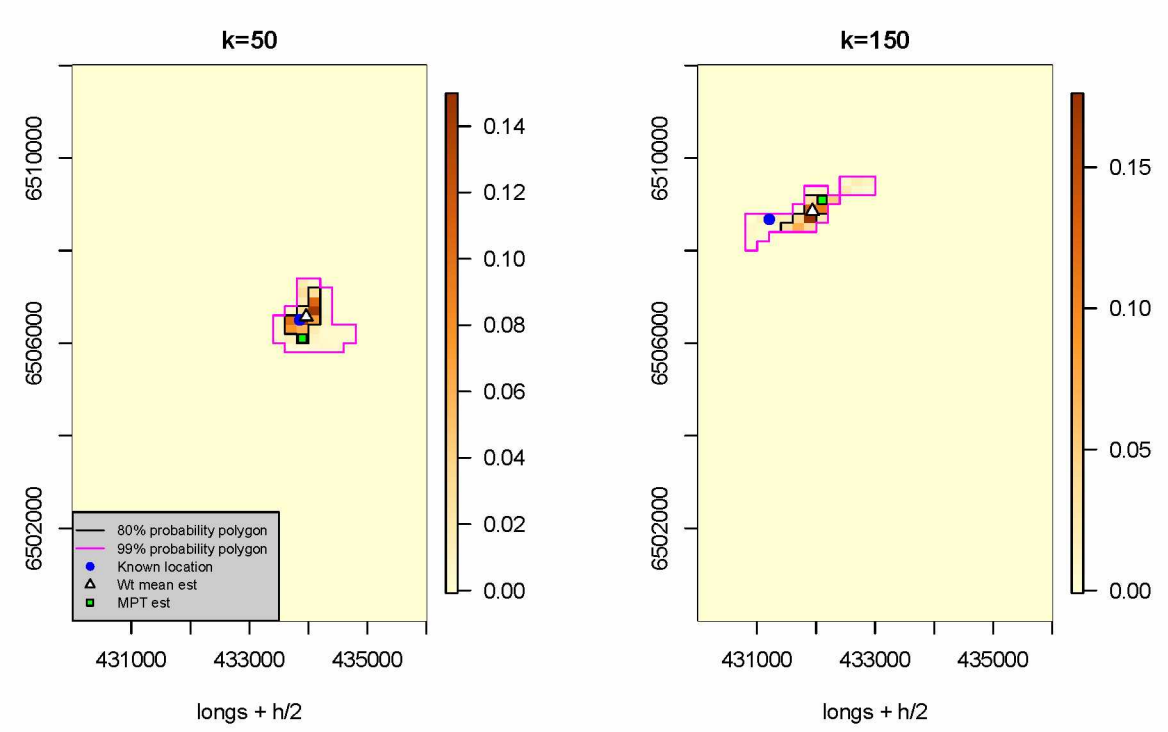


Figure A4. Examples of smoothed probability at two discrete time steps (k) for a simulated track with 200 steps on a 200-m grid (simulated and estimated paths path shown in Figure A3). Colored surface represents the probability that the simulated fish occupied each grid cell at each time step (darker colors have higher probability). The true (simulated) position at the time step is shown by a solid blue circle, the location estimated by the most probable track by a green square, and the location estimated by the weighted mean probability by a yellow triangle. Error in the weighted mean estimate is shown by polygons that encompass 80% (black line) and 99% (magenta line) of the probability surface at each time step.

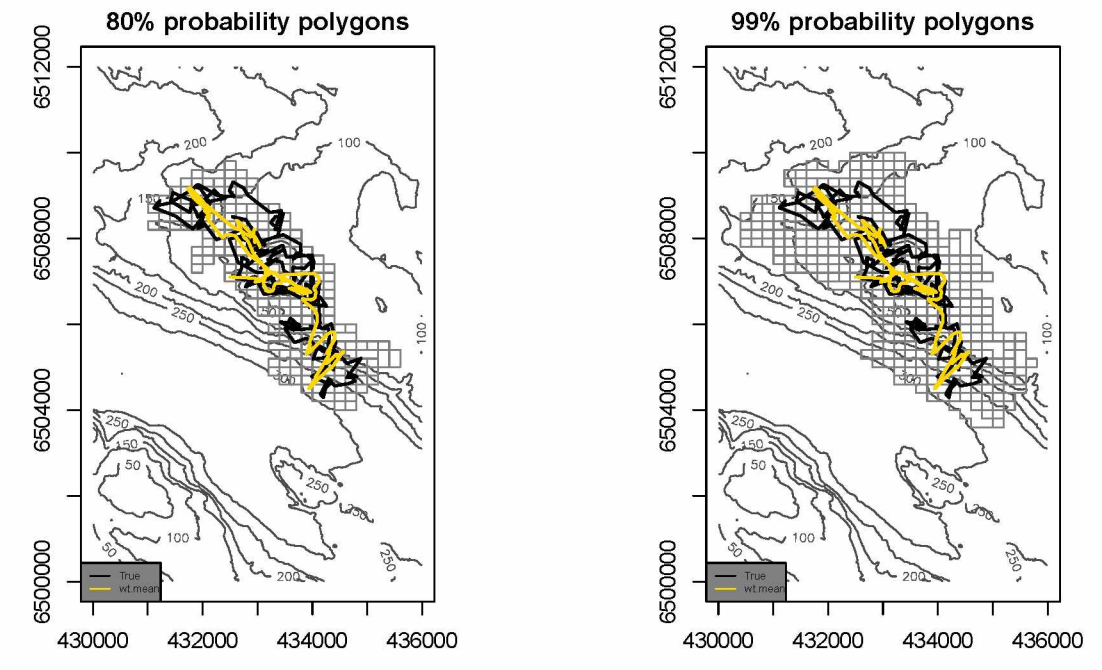


Figure A5. Example of 80% (left) and 99% (right) probability polygons for all 200 steps of a simulated trajectory (black line) estimated on a 200-m grid. Weighted mean reconstructed path shown in yellow.